

Differential Geometry: AN AUTISTIC INTRODUCTION

Alexander Bogatskiy and Umang Mehta

Current draft: May 3, 2024

Contents

I Basic Topology and Geometry	10
1 Category Theory I	10
1.1 Categories	10
1.2 Functors	11
1.3 Further examples in concrete categories	15
1.4 Classes of morphisms	20
1.5 Universal constructions	23
1.5.1 Universal objects	23
1.5.2 Universal properties	24
1.5.3 Adjoint functors	28
1.5.4 Universal constructions as adjoints	30
1.6 Products and coproducts	31
1.7 Natural transformations and category equivalence	35
2 General Topology	42
2.1 Topological spaces	42
2.1.1 Open and closed sets	42
2.1.2 Generated topologies	46
2.1.3 Connectedness	48
2.1.4 Compactness	50
2.2 Metric spaces	51

2.3	Separation axioms	54
2.4	Homotopy	55
2.5	The fundamental group	57
3	Manifolds	69
3.1	Topological manifolds	69
3.2	Smooth structures	71
3.3	Partitions of unity	73
3.4	Sard's theorem	75
3.5	Inverse & Implicit Mapping Theorems	77
3.6	Tangent spaces	77
3.7	Rank	79
3.8	Immersions and embeddings	82
3.9	Submanifolds	84
3.10	Submersions	87
4	Covering spaces	90
4.1	Homotopy lifting	90
4.2	Morphisms of covering spaces	96
4.3	Deck transformations	99
4.4	Classification of covering spaces	105
4.5	Examples of covering spaces	111
5	Homotopy Theory	117
5.1	Homotopy groups	117
5.2	Exact sequence of homotopy groups	118
5.3	<i>CW</i> -complexes	120
5.4	Homotopy extension	127
5.5	Whitehead's theorem	131
5.6	Cellular approximation	133
5.7	<i>CW</i> approximation	135
5.8	Hurewicz fibrations	140
5.9	Serre fibrations	144
5.10	Higher connectivity	150
5.11	Excision for homotopy groups	153

II Differential Geometry	159
6 Lie Theory I: Basics	159
6.1 Lie groups	159
6.2 Topology of Lie groups	162
6.3 Lie subgroups	167
6.4 Classical Lie groups	169
6.5 CW-structure of classical groups	175
6.6 Lie group actions I	178
6.7 Semidirect products	180
6.8 Representations	182
6.9 Lie algebras: Construction I	184
6.10 Lie algebras: Construction II	187
6.11 Lie subalgebras and ideals	194
6.12 Low-dimensional Lie algebras	197
6.13 (*) Infinite-dimensional Lie groups	200
7 Fiber Bundles	209
7.1 Bundles	209
7.2 Structure groups	212
7.3 Fiber morphisms	216
7.4 Bundle morphisms	221
7.5 Flat bundles	226
7.6 Subbundles, pullbacks	230
7.7 Homotopy lifting property	234
8 Vector Bundles	242
8.1 Vector bundles	242
8.2 Tangent bundle	250
8.3 Tensor bundles	256
8.4 Whitney approximation and embedding	262
8.5 Transversality	264
8.6 K-theory	268
9 Vector Fields and Lie Derivatives	274
9.1 Vector fields and flows	274
9.2 Applications of flows	277
9.3 Vector fields as differential operators	281
9.4 Geometric approach to Lie derivatives	286

9.5	Frobenius theorem I	290
10	Differential Forms and Integration	296
10.1	Exterior algebra	296
10.2	Differential forms	299
10.3	Orientation	303
10.4	Integration	306
10.5	Vector bundle-valued differential forms	309
10.6	Frobenius theorem II	311
11	Lie Theory II: Geometric Structure	318
11.1	Invariant vector fields	318
11.2	Invariant differential forms	321
11.3	Existence of homomorphisms	330
11.4	One-parameter subgroups	339
11.5	Exponential map	342
11.6	Differential of exp	347
11.7	Normal subgroups	351
11.8	Covering groups	353
11.9	(*) Technical lemmas	355
12	Lie Theory III: Analytic Structure	361
12.1	Product in logarithmic coordinates	361
12.2	Dynkin's formula	363
12.3	Geometric BCH formula	365
12.4	Closed subgroup theorem	367
12.5	Lie's fundamental theorems	371
12.6	Analytic subgroup theorem	377
12.7	Yamabe's Theorem	383
13	Group Actions on Manifolds	386
13.1	Homogeneous spaces	386
13.2	Quotients by normal subgroups	390
13.3	Lie group actions II	393
13.4	Killing vector fields	396
13.5	Proper actions	401
13.6	Equivariant tubular neighborhoods	403
13.7	Free proper actions	406
13.8	Example: flag manifolds	409

13.9	Example: projective, conformal, and CR spheres	415
13.10	(*) Diffeomorphism groups	417
III	Structured Geometry	425
14	Principal and Associated Bundles	425
14.1	Principal fiber bundles	425
14.2	Associated Bundles	428
14.3	Principal bundle morphisms	438
14.4	Bundle reduction	442
15	Klein Geometry	448
15.1	Klein geometry	448
15.2	Locally Klein geometry	450
15.3	Homogeneous bundles	454
15.4	Frobenius reciprocity	458
15.5	Local gauges	462
16	Connections	465
16.1	General connections	465
16.2	Horizontal forms	474
16.3	Principal connections	477
16.4	Induced connections	490
16.5	Holonomy and reduction	496
16.6	Gauge transformations	506
16.7	Covariant exterior derivative	511
16.8	Koszul (linear) connections	517
16.9	Parallel transport	524
17	Cartan geometry	528
17.1	Affine connections	528
17.2	Geodesics	536
17.3	Local formulas	542
17.4	Connections on G -structures	546
17.5	Invariant principal connections	559
17.6	Invariant linear connections	559
17.7	Cartan connections	559

18 Riemannian Geometry	560
18.1 Riemannian manifolds	560
18.2 Properties of geodesics	566
18.3 Riemannian connections	570
18.4 Morse theory	570
18.5 Geometry of Lie groups	570
18.6 Symmetric spaces	570
18.7 Maximal compact subgroups	570
18.8 Bott periodicity	571
19 Complex Geometry	572
19.1 Complex and Hermitian structures	572
19.2 Complex manifolds	573
19.3 Complex differential forms	575
19.4 Maurer-Cartan equation	577
20 Symplectic geometry	577
IV Homology and Cohomology	578
21 Category Theory II	578
21.1 Limits and colimits	578
21.2 Sub-objects and quotient objects	583
21.3 Abelian categories	584
22 Homological Algebra I: Exactness	588
22.1 Exact sequences and functors	588
22.2 Direct limits and exactness	593
22.3 Diagram chasing lemmas	595
23 Cohomologies of Differential Forms	601
23.1 De Rham cohomology	601
23.2 Mayer-Vietoris sequence	603
23.3 De Rham cohomology with compact support	606
23.4 Finite dimensionality of de Rham cohomology	609
23.5 Hodge theory	610
23.6 Cohomology of Lie groups	610
23.7 Dolbeault cohomology	610

24 Homology of Topological Spaces	614
24.1 Simplicial homology	614
24.2 Homological algebra II: complexes	618
24.3 Singular homology	626
24.4 Homotopy invariance	628
24.5 Relative homology and excision	630
24.6 Local homology, orientation	636
24.7 Degree	637
24.8 Hurewicz theorem	638
24.9 Singular vs. simplicial homology	640
24.10 Axioms for homology	642
24.11 Cellular homology	643
24.12 Euler characteristic	650
25 Homological Algebra III: Derived Functors	655
25.1 Injective and projective resolutions	655
25.2 Left derived covariant functors	659
25.3 Left derived contravariant functors	663
25.4 Right derived functors	664
25.5 Tor functors	665
25.6 Ext functors	670
26 Homological Algebra IV: Applications	672
26.1 Extensions	672
26.2 (*) Higher extensions	676
26.3 Elementary examples	677
26.4 Torsion submodules	680
26.5 (*) Application in class field theory	682
26.6 Homology with coefficients	688
26.7 Künneth formula	691
27 Homological Algebra V: Spectral Sequences	691
28 Cohomology of Topological Spaces	691
28.1 Eilenberg-Zilber maps	691
28.2 Singular cohomology	691
28.3 de Rham theorem	691
28.4 de Rham currents	691
28.5 Borsuk-Ulam theorem	691

28.6	Fiber sequences	692
28.7	Simplicial sets, classifying spaces	692
28.8	Eilenberg-MacLane spaces	692
28.9	Postnikov, Whitehead towers	692
28.10	Representability of cohomology	692
29	Group (Co)Homology	692
29.1	Definition	692
29.2	Cyclic groups	695
29.3	Two resolutions	696
29.4	H_1 and H^1	698
29.5	H_2 and H^2	701
29.6	Relation to topology	703
30	Lie algebra (Co)Homology	705
31	Sheaf Cohomology	706
31.1	Presheaves	706
31.2	Sheaves	706
31.3	Resolutions of sheaves	706
31.4	Sheaf cohomology	706
32	Čech Cohomology	706
32.1	Orientability and cohomology	706
32.2	Čech cohomology and its applications	706
32.3	Čech vs. sheaf vs. de Rham cohomology	706
32.4	Čech vs. singular cohomology	706
33	Applications	706
33.1	Poincaré duality	706
33.2	Thom isomorphism	706
33.3	Intersection theory	706
33.4	Thom class and degrees	706
33.5	Euler class and Lefschetz numbers	706

V Lie Theory	707
VI Gauge Theory and Index Theorems	708
34 Fiber Sequences	708
35 Classification of Principal Fiber Bundles	708
36 Characteristic Classes	708
37 Chern-Weil Theory	708
38 Clifford Algebras and Dirac Operators	708
39 Index Theorems	708
References	709
Index	712

Part I

Basic Topology and Geometry

1 Category Theory I

1.1 Categories

Definition 1.1.1 (Categories). The language of category theory is extremely useful throughout all modern mathematics, but historically it came out of developments in algebraic topology and geometry. Here we try to give a short but intuitive introduction that is focused on “unlearning” certain bad habits from traditional mathematics and pushing the reader towards this higher-level way of thinking, which will come in very handy for us later in geometry and group theory. For a more comprehensive introduction for beginners see for example [29, 30]. A category C consists of

1. A class $\text{Ob}(C)$ of *objects*;
2. For any $X, Y \in \text{Ob}(C)$ there is a set $\text{Mor}_C(X, Y)$ (also denoted by $\text{Hom}_C(X, Y)$) of *morphisms* (arrows) from X to Y with the following properties:
 - (a) The set of morphisms is closed under *composition*: $\forall f \in \text{Mor}(A, B)$ and $\forall g \in \text{Mor}(B, C)$, $\exists h = fg = g \circ f \in \text{Mor}(A, C)$.
 - (b) The composition is associative, $(fg)h = f(gh)$.
 - (c) There is an identity morphism for every object: $\forall A \in \text{Ob}(C) \exists \text{id}_A \in \text{Mor}(A, A)$ such that for any morphisms f, g with appropriate domains and codomains, $\text{id}_A f = f$ and $g \text{id}_A = g$.

The set $\text{Mor}(A, A) = \text{Hom}(A, A)$ is also denoted $\text{End}(A, A)$ or just $\text{End}(A)$, its elements are called *endomorphisms* of A .

Notice how the notation fg for the composition where f acts first and g second is more natural for category theory than it is for set theory, because we usually don’t refer to arguments of functions here, which means that we’ll never have to write strange-looking formulas like $(fg)(x) = g(f(x))$. In the cases where we do need arguments, we will use the set-theoretical notation $g \circ f$. We will always use the set-theoretical notation outside of this section on categories.

A category can be visualized as a directed graph with the vertices being the objects of the category and the arrows being the morphisms.

Example 1.1.2. 1. A category with no objects is called the empty, or trivial, category:
 $C = \emptyset$.

2. A category with only identity morphisms is called *discrete*.
3. A category with only one object is called *monoidal*.

Example 1.1.3. The most basic example of a large category is the category Set , whose objects are all sets and the morphisms are all maps between sets, i.e. $\text{Mor}(A, B) = \text{Map}(A, B)$. One *subcategory* of Set is FinSet , i.e. the category of finite sets. FinSet is a *full subcategory* of Set , meaning that $\text{Mor}_{\text{FinSet}}(A, B) = \text{Mor}_{\text{Set}}(A, B)$. In other words, a full subcategory is a subcategory in which only the class of objects is truncated and not the sets of morphisms.

Example 1.1.4. One can truncate the set of morphisms in Set to obtain the categories Bij , Surj , Inj , where the morphisms are now bijections, surjections and injections, respectively.

Example 1.1.5. We can also consider Set_* , the category of “pointed sets”. The objects of the category are doubles (A, x) where A is a set and $x \in A$ is an element of the set (a “basepoint”). The morphisms $\text{Mor}((A, x), (B, y))$ are restricted to those maps $f : A \rightarrow B$ such that $f(x) = y$, i.e., maps that preserve basepoints.

Example 1.1.6. A generalization of Set_* is the category SetPair whose objects are pairs (A, B) where A is a set and $B \subset A$. A morphism $f : (A, B) \rightarrow (C, D)$ must preserve this structure in the sense that $f(B) \subset D$.

Example 1.1.7 (Dual Category). For any category C one can define its *dual category* C^* by

$$\begin{aligned}\text{Ob}(C^*) &= \text{Ob}(C), \\ \text{Mor}_{C^*}(A, B) &= \text{Mor}_C(B, A).\end{aligned}\tag{1.1.1}$$

In most cases, duals of categories are very difficult to interpret in any other way than directed graphs. For instance, the dual of Set has $\text{Mor}_{\text{Set}^*}(\{1, 2\}, \{1\}) = \text{Map}(\{1\}, \{1, 2\}) = \{f_1, f_2\}$, but there aren’t two maps from $\{1, 2\}$ to $\{1\}$! Thus the arrows in Set^* cannot be interpreted as maps between sets.

1.2 Functors

We can now construct “maps” from categories to categories. Crucially, we want these maps to preserve as much of the categorical structure as possible, which is why they have to act not only on objects, but also on morphisms, in a consistent way.

Definition 1.2.1 (Functors). A functor $F : C \rightarrow \mathcal{D}$ is a correspondence between two categories C and \mathcal{D} with the following properties

1. $\forall A \in \text{Ob}(C), F(A) \in \text{Ob}(\mathcal{D})$.
2. F maps objects to objects, *and* morphisms between objects to morphisms between their

images:

$$\forall f \in \text{Mor}_C(A, B), \quad F(f) \in \begin{cases} \text{Mor}_{\mathcal{D}}(F(A), F(B)), & \text{if covariant functor,} \\ \text{Mor}_{\mathcal{D}}(F(B), F(A)), & \text{if contravariant functor.} \end{cases} \quad (1.2.1)$$

3. F respects compositions: $\forall A \in \text{Ob}(C)$ we have $F(\text{id}_A) = \text{id}_{F(A)}$ and

$$F(fg) = \begin{cases} F(f)F(g), & \text{if covariant,} \\ F(g)F(f), & \text{if contravariant.} \end{cases} \quad (1.2.2)$$

Covariant functors can be thought of as functors that preserve the direction of arrows between objects while contravariant functors reverse the arrows.

The definition of a (covariant) functor is set up in such a way that any sequence of morphisms $A_1 \xrightarrow{f_1} A_2 \rightarrow \dots \xrightarrow{f_n} A_{n+1}$ (or simply $f_1 f_2 \dots f_n$) in C is mapped to exactly one morphism $F(f_1)F(f_2)\dots F(f_n)$ in \mathcal{D} . This seemingly trivial property has incredibly deep implications. Represented as a diagram for the case $n = 2$, this property looks like this:

$$\begin{array}{ccc} F(A_1) & \xrightarrow{F(f_1 f_2)} & F(A_3) \\ & \searrow F(f_1) \curvearrowright \nearrow F(f_2) & \\ & F(A_2) & \end{array} \quad (1.2.3)$$

The circular arrow here indicates that the composition of morphism along any path in this triangle must give the same morphism, i.e. precisely that $F(f_1 f_2) = F(f_1)F(f_2)$. When this holds, we say that “the diagram commutes”. Now suppose $A_3 = A_1$ and $f_1 = f_2^{-1}$ is an isomorphism between A_1 and A_2 . Then for any functor F , not only is it true that $F(\text{id}_{A_1})$ is still an identity morphism, but also that $F(f_1)$ is still an isomorphism, since $F(f_1)F(f_2) = \text{id}_{F(A_1)}$ and $F(f_2)F(f_1) = \text{id}_{F(A_2)}$. Intuitively, this means that *the definition of F itself cannot use any information contained in the objects of C that is not preserved by isomorphisms*.

Example 1.2.2 (Non-examples of functors). Some common constructions are not functors. For example, for a set X one can define the group $\text{Aut}(X)$ of bijections (permutations) $X \rightarrow X$. This is not a functor because there is no nontrivial way to define its action on maps (the trivial way would be to send everything to the identity map). Similarly, the monoid (same as a group but without inversion) $\text{End}(X)$ of endomorphisms (i.e. all maps $X \rightarrow X$) is also not functorial in X .

The word “functorial” is a general adjective that expresses the idea that many useful constructions in mathematics respect morphisms. Very roughly, this “principle of naturality” can be stated as:

Isomorphic inputs must produce isomorphic outputs.
In particular, all constructions should not give preference to any
specific representatives of isomorphism classes,
i.e. they must be *natural*.

(1.2.4)

Looking ahead, an even deeper generalization of this principle will arise when we go up to the level of *morphisms between functors* (called natural transformations).

Example 1.2.3 (Hom-functors). One can always construct a functor from any category C to Set , as in the following example: for a fixed object $A \in \text{Ob}(C)$ consider the quantity

$$h^A(_) = \text{Mor}_C(A, _), \quad (1.2.5)$$

where the $_$ is where the argument of h^A goes, i.e., $h_A(B) = \text{Mor}_C(A, B)$. Since the morphisms between any two objects by definition form a set, h^A is a map from $\text{Ob}(C)$ to $\text{Ob}(\text{Set})$. Can we extend h^A to a functor by defining a correspondence between morphisms? The answer is yes, and here's how.

Consider a morphism $f : B \rightarrow C$ in the category C . We require correspondingly a map $h^A(f) : h^A(B) \rightarrow h^A(C)$. $h^A(B)$ is the set of maps from A to B and $h^A(C)$ is the set of maps from A to C . Hence, what we require is the following: Given a map from B to C , how do we transform a map $g : A \rightarrow B$ into a map from A to C ? The answer is composition, of course! Hence the map $h^A(f)$ acts on g by the action

$$h^A(f)(g) = gf = f \circ g. \quad (1.2.6)$$

$$\begin{array}{ccc} B & \xrightarrow{f} & C \\ g \swarrow & & \nearrow gf \\ A & & \end{array} \quad (1.2.7)$$

One can check that h^A is now a *covariant* functor (Exercise).

Similarly, we can also construct the *contravariant hom-functor* by defining

$$h_A(_) = \text{Mor}_C(_, A). \quad (1.2.8)$$

In this case, its action on morphisms will be achieved by composition in the opposite direction. These functors are called *representable functors*, or *hom-functors*.

Example 1.2.4. Any partially ordered set (“poset”) (I, \leq) can be represented by a directed graph where we draw an arrow from $i \in I$ to $j \in I$ if and only if $i \leq j$. By reflexivity and transitivity of inequality, this satisfies the definition of a category. Therefore each poset (I, \leq) is also a category \mathcal{I} . Functors between posets are exactly the *monotonic* maps.

Definition 1.2.5 (Arrow category). For any category C we can construct its arrow category $\text{Arr}(C)$ as follows:

1. Objects are all arrows in C :

$$\text{Ob}(\text{Arr}(C)) = \bigcup_{A, B \in \text{Ob}(C)} \text{Mor}_C(A, B). \quad (1.2.9)$$

2. Morphisms between arrows are pairs of morphisms between their ends that form *commuting squares*:

$$\begin{array}{ccc} A_1 & \xrightarrow{f_1} & B_1 \\ \alpha \downarrow & \curvearrowright & \downarrow \beta \\ A_2 & \xrightarrow{f_2} & B_2 \end{array} \quad (1.2.10)$$

i.e. $f_1\beta = \alpha f_2$. The circular arrow is used to indicate that the square commutes.

Exercise 1.1. Check that this is a category.

Definition 1.2.6 (Categories of diagrams). We can generalize the definition of arrow categories to define *categories $C^{\mathcal{J}}$ of \mathcal{J} -shaped diagrams in C* . First, let \mathcal{J} be a so-called *index category*, which is an arbitrary, usually small (in most cases even finite), category. It should be thought of as nothing but a directed graph. A *\mathcal{J} -shaped diagram* in C is defined as a covariant functor $F : \mathcal{J} \rightarrow C$. Note that we don’t require this functor to be faithful (injective on objects), so some of the “vertices” of \mathcal{J} might map to the same object in C . Now, the category of \mathcal{J} -shaped diagrams in C has all \mathcal{J} -shaped diagrams as its objects, and morphisms between two \mathcal{J} -shaped diagrams F_1 and F_2 are sets of pairwise morphisms $\alpha_i : F_1(i) \rightarrow F_2(i)$, $i \in \text{Ob}(\mathcal{J})$, such that all “vertical” squares commute: for $f_1 : i \rightarrow j$ and $f_2 : i \rightarrow j$, we have a commuting square

$$\begin{array}{ccc} F_1(i) & \xrightarrow{F_1(f_1)} & F_1(j) \\ \alpha_i \downarrow & \curvearrowright & \downarrow \alpha_j \\ F_2(i) & \xrightarrow{F_2(f_2)} & F_2(j) \end{array} \quad (1.2.11)$$

A morphism in this category can be visualized as a prism with the top and bottom sides being two copies of \mathcal{J} , all vertical arrows in the same direction, and all side squares commuting.

This construction is also called a *natural transformation* between the functors F_1 and F_2 (written as $F_1 \Rightarrow F_2$) and is a higher-level generalization of morphisms (acting on objects) and functors (acting on morphisms), now acting on functors.

Of course, if the index category is just one arrow, $\mathcal{J} = \{1 \rightarrow 2\}$ (identity morphisms are implicit), then we get back the arrow category $\text{Arr}(C)$.

Example 1.2.7. A yet further generalization of the above construction is one where some vertices and/or morphisms in $F(\mathcal{J})$ are fixed, which would be a full subcategory of $\mathcal{J}\text{-Diag}(C)$. Moreover, in concrete categories some of the fixed vertices may even belong to a different category than C , with morphisms attached to them being a special kind of maps between objects of different categories (e.g. maps from a set into a group, which is an example we will need when defining free groups).

Example 1.2.8. Let's consider some slightly more complicated examples of categories, starting with Pow , the power set category.¹ The objects of this category are sets, as usual. But the morphisms are defined as $\text{Mor}_{\text{Pow}}(X, Y) = \text{Map}(2^X, 2^Y)$.

One can always construct a map P that takes a set X to its power set 2^X . How can this map be extended to a functor from Set to itself? In other words, how can we define, for a morphism $f : X \rightarrow Y$, a morphism $P(f) : 2^X \rightarrow 2^Y$ for a covariant functor or $2^Y \rightarrow 2^X$ for a contravariant functor? The answer is simple. For the covariant case, pick a subset $A \subset X$ and define

$$P_+(f)(A \subset X) = f(A), \quad (1.2.12)$$

i.e., the image of A in Y . It is easy to check the *functoriality* property $P_+(fg) = P_+(f)P_+(g)$. For the contravariant case we can similarly define

$$P_-(f)(B \subset Y) = f^{-1}(B), \quad (1.2.13)$$

where $f^{-1}(B)$ is the pre-image of B in X , i.e., the set of elements of X that map into B under f . The two functors above are called the *power functors*.

1.3 Further examples in concrete categories

Concrete categories are categories whose objects are sets with some additional structure and whose morphisms are maps respecting that structure. For instance, the category Gr of groups² with the morphisms being group homomorphisms.

Example 1.3.1 (Commutant and abelianization functors). For any group G , we can define its *commutant* subgroup $[G, G]$ as the group of elements of the form $\{ghg^{-1}h^{-1} \mid g, h \in G\}$. Correspondingly, we can define a functor $F : \text{Gr} \rightarrow \text{Gr}$ that takes every group G to its

²Recall that groups are sets with a closed, associative, invertible multiplication and an identity element.

commutant subgroup $[G, G]$. The functor acts on homomorphisms by restricting them to their action on the commutant subgroup, i.e., for $f : G \rightarrow H$, $F(f) = f|_{[G, G]}$.

Commutant subgroups are also normal subgroups³ and one can verify that the quotient⁴ $G/[G, G]$ is an abelian group, called the *abelianization* of G . Hence, one can define the abelianization functor $F : \text{Gr} \rightarrow \text{Ab}$ from the category of groups to the category of abelian groups (the morphisms are still group homomorphisms). For a group homomorphism f in Gr , the homomorphism $F(f)$ can be defined by considering its action on any representative element of an equivalence class (under the conjugation relation) in G . The abelianization functor will be useful later when we study the relation between homotopy and homology groups of a topological space.

Example 1.3.2 (Quotient functor). Another example from category theory is the category GrPair of pairs (G, H) , where G is a group and H is its normal subgroup, $H \trianglelefteq G$. Morphisms are group homomorphisms $f : G_1 \rightarrow G_2$ such that the image of H_1 is contained in H_2 , $f(H_1) \leqslant H_2$ (inequality sign signifies that the l.h.s. is also a subgroup in the r.h.s.). We can construct the *quotient functor* $F : \text{GrPair} \rightarrow \text{Gr}$ such that $F((G, H)) = G/H$ and $F(f) : G_1/H_1 \rightarrow G_2/H_2$ is the induced homomorphism of the quotients. The existence and uniqueness (i.e. naturality) of this homomorphism (and that it preserves compositions) is ensured by the **fundamental homomorphism theorem**. Therefore the fundamental homomorphism theorem is nothing but the statement that F is naturally a functor. Note that by construction, $H_1 \leqslant \ker F(f)$. If $H_1 = \ker F(f)$, then $H_2 = \{e\}$ and $F(f) : G_1/\ker f \rightarrow G_2$ must be an isomorphism by uniqueness, so we get back the **first isomorphism theorem** that $f(G_1) = G_2 \cong G_1/\ker f$.

Example 1.3.3. One more useful example of a category is the category C of all maps between sets and groups (both of these can be replaced by other concrete categories):

1. $\text{Ob}(C) = \bigcup_{X, G} \text{Map}(X, G)$, where the union is over all $X \in \text{Ob}(\text{Set})$ and $G \in \text{Ob}(\text{Gr})$.
2. $\text{Mor}\left(X \xrightarrow{g} G, Y \xrightarrow{h} H\right)$ consists of pairs of morphisms, $X \xrightarrow{\alpha} Y$ in Set , and $G \xrightarrow{\beta} H$ in Gr , which make the following square commute:

$$\begin{array}{ccc} X & \xrightarrow{\alpha} & Y \\ g \downarrow & \curvearrowright & \downarrow h \\ G & \xrightarrow{\beta} & H \end{array} \quad (1.3.1)$$

Exercise 1.2. Come up with more examples of categories following the structure of the last example, including where morphisms are not necessarily squares (triangles, prisms,...).

Definition 1.3.4 (Rings). A ring $(R, +, \cdot)$ is defined as a set R with two binary operations $+$ and \cdot such that

1. $(R, +)$ is an abelian group.
2. \cdot is distributive over $+$, i.e. $a \cdot (b + c) = a \cdot b + a \cdot c$.
3. (If the ring is associative) \cdot is associative.
4. (If the ring is associative and has a unity) there is a multiplicative neutral element denoted by 1 , and $1 \neq 0$, so that $R \neq \{0\}$.

Some examples of rings include the ring of integers $(\mathbb{Z}, +, \cdot)$, and the ring $\mathbb{R}[x]$ of polynomials in a variable x with real coefficients. Note that requirement associativity of multiplication is sometimes relaxed in the definition of a ring.

We can define two categories:

- **Ring** of all rings, associative or not, with morphisms being ring homomorphisms (maps that preserve addition, multiplication).
- **AssRing** of associative rings with 1 , the multiplicative identity, such that $1 \neq 0$, and the morphisms being ring homomorphisms (preserving the unity).
- **CommRing** of associative, commutative rings with $1 \neq 0$.

Example 1.3.5. Since every ring is an abelian group with respect to $+$ by definition, we can define a *forgetful functor* $F : \text{AssRing} \rightarrow \text{Ab}$ that maps every ring to its corresponding abelian group by 'forgetting' the multiplication structure of the ring.

Example 1.3.6. We can also define a functor $F : \text{Ring} \rightarrow \text{Ring}$ that maps every ring R to the ring $R[x]$ defined as the ring of polynomials in x whose coefficients are taken from R . Then for a morphism f acting on elements of the ring R , $F(f)$ simply applies f to every coefficient of the polynomial.

Example 1.3.7. Just as we can define square matrices of real numbers, we can also define square matrices with coefficients taken from a commutative ring R , the set of matrices being denoted by $\text{Mat}(n, R)$, n being the size of the matrix. $\text{Mat}(n, R)$ forms an associative ring. Hence, we have a covariant functor $F : \text{CommRing} \rightarrow \text{AssRing}$ that takes a commutative ring to matrices of the ring.

Exercise 1.3. How does this functor act on morphisms? Check its functoriality.

Example 1.3.8. If we restrict to invertible matrices the associative ring $\text{Mat}(n, R)$ reduces to the group $\text{GL}(n, R)$ of invertible matrices with coefficients in R and we have a functor $F : \text{CommRing} \rightarrow \text{Gr}$.

Example 1.3.9. For every associative ring R we can define the *opposite ring* R^{op} which has the same elements and addition as R , but the “opposite” multiplication defined by $a \cdot^{\text{op}} b = b \cdot^R a$, where \cdot^R is the multiplication from R . This is a contravariant functor $\text{AssRing} \rightarrow \text{AssRing}$.

Definition 1.3.10 (Lie rings). A Lie ring is a set R with a binary operation $[\cdot, \cdot]$ (called the Lie bracket) that obeys the following properties $\forall a, b, c \in R$

1. Alternativity: $[a, a] = 0$;
2. Antisymmetry: $[a, b] = -[b, a]$;
3. Jacobi identity: $[a, [b, c]] + \text{cyclic permutations} = 0$.⁵

The category LieRing consists of all Lie rings as objects with Lie ring homomorphisms (maps that preserve the Lie bracket) as the morphisms.

Example 1.3.11. For any ring R we can define an associated Lie ring $R^{(-)}$ by defining $[a, b] = a \cdot b - b \cdot a$ and ignoring the multiplication and addition structure after defining the Lie bracket. One can check that this obeys the properties of a Lie ring. This defines a functor $F : \text{AssRing} \rightarrow \text{LieRing}$.

Definition 1.3.12 (Fields). A field K is a set with two binary operations $+$ (addition) and \cdot (multiplication) such that

1. $(K, +, \cdot)$ is a *commutative* ring;
2. Every element in $K^* = K \setminus \{0\}$ has a multiplicative inverse (0 being the additive identity), i.e. K^* is an abelian group.

A morphism between fields is a map that preserves addition, multiplication, zero and unity. The category of fields is rarely referred to because of its much weaker properties than that of Ring and others (it is not “algebraic” in some strict sense).

K^* is called the multiplicative group of the field.

Definition 1.3.13 (Ring modules). A (left) module M over a ring R (also called a left R -module) is an abelian group $(M, +)$ with an operation $\bullet : R \times M \rightarrow M$ such that $\forall r, s \in R$ and $x, y \in M$ the following properties hold

1. $r \bullet (x + y) = r \bullet x + r \bullet y$;
2. $(r + s) \bullet x = r \bullet x + s \bullet x$;
3. $(r \cdot s) \bullet x = r \bullet (s \bullet x)$;
4. (If the ring is unital, i.e., consists of a unit multiplicative element) $1_R \bullet x = x$.

Similarly one can define *right* R -modules where elements of the ring act from the right, $x \bullet r$. Note that a left R -module is exactly the same as a right R^{op} -module, so for general constructions and arguments left modules suffice.

An R_1 - R_2 -*bimodule* is a left R_1 -module and a right R_2 -module at the same time, with the compatibility requirement $(r_1 \bullet x) \bullet r_2 = r_1 \bullet (x \bullet r_2)$ (denoting both products by \bullet because confusion is not possible here).

A *morphism* of (left) R -modules is a group homomorphism $f : M \rightarrow L$ that respects the action of the ring, i.e. $f(r \bullet_M x) = r \bullet_L f(x)$.

The categories of left and right R -modules are denoted by $R\text{-Mod}$ and $\text{Mod-}R$, respectively.

The symbol for ring multiplication is often omitted in order to simplify notation and avoid confusion with the scalar multiplication operator \bullet .

Example 1.3.14. For any ring R , its Cartesian powers R^n are called *free R -modules*. The action of R on it is defined component-wise: $r \cdot (r_1, \dots, r_n) = (rr_1, \dots, rr_n)$.

Example 1.3.15. R^n is an example of a left $\text{Mat}(n, R)$ -module, where the action is defined by the usual matrix multiplication.

Vector spaces are nothing but a special case of ring modules, where the ring is replaced by a field.

Definition 1.3.16 (Vector spaces). A vector space V over a field K is simply a K -module, i.e. an abelian group with a multiplicative action of the field K . Morphisms between vector spaces are therefore K -module morphisms (a.k.a. “linear maps”). The category of K -vector spaces is Vect_K . The category of finite-dimensional K -vector spaces is FinVect_K .

Exercise 1.4. Show that abelian groups are essentially the same as \mathbb{Z} -modules.

Definition 1.3.17 (Algebras over fields). An algebra \mathcal{A} over a field K is a K -vector space with a binary, K -bilinear multiplication operation \cdot . In other words, $\forall x, y, z \in \mathcal{A}$ and $\forall a, b \in K$

1. $(x + y) \cdot z = x \cdot z + y \cdot z;$
2. $z \cdot (x + y) = z \cdot x + z \cdot y;$
3. $(ax) \cdot (by) = (ab)(x \cdot y).$

Notice that we do not require all algebras to be associative (this way Lie brackets are allowed as the multiplication operation). Algebra morphisms are already defined by the above definitions. The category of K -algebras is denoted by Alg_K .

Example 1.3.18. Some examples of algebras are $\mathbb{R}[t]$, the algebra of polynomials in t with real coefficients; \mathbb{C} , the set of complex numbers viewed as an algebra over \mathbb{R} ; and $(\mathbb{R}^3, +, \times)$ with \times being the cross product for vectors⁶.

Exercise 1.5. Give the definitions of:

- division rings (non-abelian analog of fields, e.g. quaternions);
- group modules (category $G\text{-Mod}$);
- algebras over a commutative ring R (category Alg_R);
- Lie algebras (category LieAlg).

1.4 Classes of morphisms

In set theory, maps between sets can have the special properties of being surjective, injective or both, i.e., bijective. In this section we introduce the analogs of such maps in category theory. Surjective maps are those whose image spans the entire codomain, while for injective maps, every element of the range of the map has a unique pre-image. These definitions can't be generalized to categories, since the objects of a category may not have any internal structure similar to that of a set. However, there is an alternate way to define surjective and injective maps that does not rely upon the internal structure of sets, and is equivalent to the previous definitions.

$$\text{"Epic" property: } X \xrightarrow{f} Y \begin{array}{c} \nearrow g \\ \searrow h \end{array} Z \quad (1.4.1)$$

Suppose $f : X \rightarrow Y$ is a surjective map. Consider two maps $g, h : Y \rightarrow Z$ such that $g \circ f = h \circ f$ (see figure above). Since f is surjective, g and h have to agree on all of Y , which implies that $g = h$. Conversely, if $g \circ f = h \circ f$ for any g, h implies $g = h$, we can conclude that f is surjective. This is the alternate definition of a surjective map. This property is also known as *right invertibility* or *right cancellability*.

For injective maps, the alternate definition is similar. A function $f : X \rightarrow Y$ is invertible iff $\forall g, h : Z \rightarrow X, f \circ g = f \circ h \longleftrightarrow g = h$, i.e., f is injective iff it is *left invertible* (or *left cancellable*).

The generalization of these definitions to categories is now straightforward (up to switching “left” and “right” because our categorical notation is opposite of the set theoretical one, $gf = f \circ g$).

Definition 1.4.1 (Epimorphism). A morphism f is an *epimorphism* if $fg = fh$ for any g, h implies $g = h$ (i.e. left cancellative). We use the arrow \twoheadrightarrow to denote epimorphisms.

Definition 1.4.2 (Monomorphism). A morphism f is a *monomorphism* if $gf = hf$ for any g, h implies $g = h$ (i.e. right cancellative). We use the arrow \rightarrowtail to denote monomorphisms.

Definition 1.4.3 (Bimorphism). A morphism f is a *bimorphism* if it is both an epimorphism and a monomorphism.

Definition 1.4.4 (Retraction/Split epi). A morphism $f : A \rightarrow B$ is a *retraction*, or a *split epi*, if $\exists g$ such that $gf = \text{id}_B$ (i.e. f is left invertible in categorical notation). In this case g is a *split mono*.

Definition 1.4.5 (Coretraction/Section/Split mono). A morphism $f : A \rightarrow B$ is a *coretraction*, or a *section*, or a *split mono*, if $\exists g$ such that $fg = \text{id}_A$ (i.e. f is right invertible). In this case

| g is a split epi, f is a section of g , and g is a retraction of f .

Definition 1.4.6 (Isomorphism). A morphism f is an *isomorphism* if it's a retraction and a coretraction (i.e. invertible). Sometimes we use the arrow \leftrightarrow to denote isomorphisms. If two objects A and B are related by an isomorphism f , we might write $A \xrightarrow{f} B$ or just $A \cong B$.

Note in particular that, unlike for sets, left(right) invertibility is not equivalent to left(right) cancellability. An important characteristic of a category is whether all bimorphisms are isomorphisms.

Proposition 1.4.7. In Set , monomorphism = coretraction = injectivity, and epimorphism = retraction = surjectivity.

Proposition 1.4.8. In Gr , the category of groups, a monomorphism is equivalent to an injective homomorphism and an epimorphism is identical to a surjective homomorphism.

In all concrete categories, the following implications hold

$$\text{Retraction} \implies \text{Surjectivity}, \quad \text{Surjectivity} \implies \text{Epimorphism}, \quad (1.4.2)$$

$$\text{Coretraction} \implies \text{Injectivity}, \quad \text{Injectivity} \implies \text{Monomorphism}. \quad (1.4.3)$$

while the converse don't. Here are some counterexamples.

Example 1.4.9 ($\text{Epi} \Rightarrow \text{Surj}$). Consider the inclusion map⁷ $i : \mathbb{Z} \hookrightarrow \mathbb{Q}$ in the category Ring . We will show that i is an epimorphism in the category of rings, while it is clearly not surjective. For two arbitrary ring morphisms $f, g : \mathbb{Q} \rightarrow R$, $f \circ i = g \circ i \implies f|_{\mathbb{Z}} = g|_{\mathbb{Z}}$. Since f and g are ring homomorphisms and any rational number can be written as a ratio of integers, we see that

$$f\left(\frac{m}{n}\right) = \frac{f(m)}{f(n)} = \frac{g(m)}{g(n)} = g\left(\frac{m}{n}\right). \quad (1.4.4)$$

Hence, $f = g$ for all rational numbers and i is an epimorphism. A nearly identical proof holds for the inclusion map $\mathbb{Q} \hookrightarrow \mathbb{R}$ in the category TopRing of topological rings⁸ is an epimorphism, except that in this case we use the fact that a continuous function defined on the rational numbers uniquely defines a continuous function on the real line because rationals are dense in the real line.

Example 1.4.10 ($\text{Surj} \Rightarrow \text{Split epi}$). Another example is the "mod 2" function $\mathbb{Z} \xrightarrow{\text{mod } 2} \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$, which is clearly surjective but is not a split epimorphism. In order for the map to be a split epimorphism, there must exist a morphism f from $\mathbb{Z}/2\mathbb{Z}$ to the integers such that $(\text{mod } 2) \circ f = \text{id}_{\mathbb{Z}/2\mathbb{Z}}$. However, besides the trivial morphism $f = 0$, there are no ring morphisms $\mathbb{Z}/2\mathbb{Z} \rightarrow \mathbb{Z}$ at all! The identity element 0 must be mapped to itself, so all that

remains is to pick an integer n that 1 maps into. But, in $\mathbb{Z}/2\mathbb{Z}$, $1 + 1 = 0$ while $n + n = 2n \neq 0$. Thus a section for mod 2 does not exist.

Exercise 1.6. Come up with examples that show that Mono \Rightarrow Inj and Inj \Rightarrow Split mono.

Example 1.4.11 (Arithmetic isn't trivial). Let $f : G \rightarrow K$ be a group homomorphism in Gr. By the first isomorphism theorem

$$f(G) \cong G/\ker(f) \quad (1.4.5)$$

and $H = \ker(f)$ is a normal subgroup of G . Hence we can construct the following sequence of maps

$$1 \xrightarrow{\text{mono}} H \xrightarrow{\text{mono}} G \xrightarrow{f} G/H \xrightarrow{\text{epi}} 1 \quad (1.4.6)$$

The first map is unique because 1 has to map to $e \in H$. The second map is the inclusion of the subgroup $H \hookrightarrow G$. These two maps are monomorphisms because they are injective, while the other two are epimorphisms because they are surjective. The question we seek to answer is whether the group G can always be written as a direct product of its normal subgroup H and the quotient group G/H . This statement is equivalent to the statement that there exists a morphism $g : G/H \rightarrow G$ such that $gf = \text{id}_{G/H}$, i.e. g is a section of the split epimorphism f .

We already know from the last example that not every epimorphism has a section, but for another example let us go back to the kindergarten and re-learn how to add integer numbers in their decimal representation. For simplicity we will restrict ourselves to two digit numbers, i.e. elements of the cyclic group $\mathbb{Z}/100\mathbb{Z}$. The units and tens digits of these numbers both form their own cyclic groups $\mathbb{Z}/10\mathbb{Z} = \{0, 1, \dots, 9\}$ and we have the following sequence of maps

$$0 \hookrightarrow \underbrace{\mathbb{Z}/10\mathbb{Z}}_{\text{tens}} \xrightarrow{\times 10} \mathbb{Z}/100\mathbb{Z} \xrightarrow{\text{mod } 10} \underbrace{\mathbb{Z}/10\mathbb{Z}}_{\text{units}} \rightarrow 0 \quad (1.4.7)$$

Just like in the example with $\mathbb{Z}/2\mathbb{Z}$ above, the map “mod 10” is epi but not a split epi, so $\mathbb{Z}/100\mathbb{Z}$ is not a trivial extension of $\mathbb{Z}/10\mathbb{Z}$ by itself. If we had $\mathbb{Z}/100\mathbb{Z} \cong \mathbb{Z}/10\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$, addition would be trivial and we would never have to “carry over” from the units digit to the tens digit. Clearly, that is not true and this isomorphism doesn't exist.

When f above isn't split, we say that G is a nontrivial *extension* of the group H .⁹

1.5 Universal constructions

1.5.1 Universal objects

The notion of universal constructions is one of the most common ways category theory shows up in applications. Here we only give a brief introduction to these concepts, but the reader is encouraged to check out the great book [28] for a much deeper yet very pedagogical exploration of universality.

Definition 1.5.1 (Initial and terminal objects). An object $A \in \text{Ob}(C)$ is called *initial* (*terminal*) if for any object $B \in \text{Ob}(C)$, the set $\text{Mor}(A, B)$ (respectively, $\text{Mor}(B, A)$) consists of exactly one element. It is called a *zero* (or *null*) object if it is both initial and terminal.

Proposition 1.5.2. *If A and A' are both initial (terminal, null) objects in C , then they are isomorphic via a unique isomorphism.*

Proof. We have $\exists! f : A \rightarrow A'$ and $\exists! g : A' \rightarrow A$. In addition, $\text{Mor}(A, A) = \{\text{id}_A\}$ and $\text{Mor}(A', A') = \{\text{id}_{A'}\}$. At the same time, $fg \in \text{Mor}(A, A)$, which means that $fg = \text{id}_A$. Similarly, $gf = \text{id}_{A'}$. Therefore f is an isomorphism, and it is unique. ■

In particular categories, the existence of universal objects often has to be shown by an explicit construction of one. However, the concept is still very powerful because it defines these objects up to isomorphism and, as such, distills their essential properties down to a very simple recipe.

Example 1.5.3. In Set , \emptyset is an initial object, because there is a unique map from \emptyset to any other set (this map is also denoted by $\emptyset : \emptyset \rightarrow A$). Terminal objects here are all one-element sets (“singletons”), collectively denoted by $\{\ast\}$, emphasizing that they are important only up to isomorphism.

Example 1.5.4. Following Example 1.2.4, in a category that is a *totally ordered* set (i.e. any two elements can be compared), a terminal object is a maximal element and an initial object is a minimal element (if they exist).

Example 1.5.5. Set_\bullet has a significantly different structure from Set , and this shows in its universal objects, because the empty set is no longer in this category. Thus one-element sets become zero objects!

Example 1.5.6. In Gr and Ab , the trivial one-element group is a zero object for reasons very similar to Set_\bullet .

Example 1.5.7. In CommRing, the ring \mathbb{Z} of integers is initial. There are no terminal objects here, because the trivial ring $\{0\}$ is forbidden by the requirement of the existence of unity $1 \neq 0$.

1.5.2 Universal properties

Definition 1.5.8 (Free group). The *free group* F_X generated by a set X is usually defined as the group whose elements are all the finite-length “words” formed out of elements x of X and their formal inverses x^{-1} as letters, so $F_X = \{x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_n^{\epsilon_n} | n \geq 0, \epsilon_i \in \{\pm 1\}, x_i \in X, i = 1, \dots, n\}$. The group operation is simply the concatenation of words and the inversion involves reversing the order of the letters and inverting each one. Two words are considered equivalent if one is obtained from the other by removals or insertions of any trivial sub-strings of the form $x^\epsilon x^{-\epsilon}$. We will also sometimes denote F_X by $\langle X \rangle$.

Now, this definition constructs F_X as an explicit set of elements using set-theoretical language. However, in group theory we are only interested in objects up to isomorphism. In particular, one could come up with infinitely many other ways to construct F_X by introducing extra letters, replacing them with other elements whatsoever, allowing for exponents other than ± 1 , etc. This would not change the group as far as category theory is concerned. Therefore it is important to understand how this construction, up to isomorphism, can be expressed in purely categorical terms. The following example shows how free groups can be defined as universal objects in a certain category. Such properties of special constructions are called *universal properties*.

Example 1.5.9 (Over/undercategory). This example will be a very common special case of what's called a “comma category”, known as an *undercategory* or a *coslice category*. Fix a set X and consider the “arrow category” C where objects are all maps from X to any group

- $\text{Ob}(C) = \bigcup_{G \in \text{Ob}(\text{Gr})} \text{Map}(X, G) = \{\text{all maps } \psi : X \rightarrow G \mid G - \text{any group}\}$.
- $\text{Mor}(\psi_1, \psi_2)$ consists of all commuting triangles with a *group homomorphism* at the bottom:

$$\begin{array}{ccc}
 & X & \\
 \psi_1 \swarrow & \curvearrowright & \searrow \psi_2 \\
 G_1 & \xrightarrow{f} & G_2
 \end{array} \tag{1.5.1}$$

Now consider looking for an initial object in this category. Such an object would consist of a

group G_X and a map $i : X \rightarrow G_X$ such that for any other map $\psi : X \rightarrow G$, there exists a *unique* group homomorphism $\psi_* : G_X \rightarrow G$ that makes the following triangle commute:

$$\begin{array}{ccc}
 & X & \\
 i \swarrow & \circlearrowleft & \searrow \psi \\
 G_X & \dashrightarrow_{\exists! \psi_*} & G
 \end{array} \tag{1.5.2}$$

In other words, any group-valued function ψ on X “factors” through G_X and the universal map i as $\psi = i\psi_*$.

We can now easily verify that such an object exists and it is isomorphic (in C) to the object $X \xrightarrow{i} F_X$ where $i(x) = "x"$. Indeed, if any G -valued map ψ factors through F_X via a homomorphism $\psi_* : F_X \rightarrow G$, then by commutativity of the triangle, $\psi_*(x) = \psi_*(i(x)) = \psi(x)$. However, one-letter words generate the whole F_X , so this equality *uniquely* defines ψ_* as a homomorphism on F_X , namely

$$\psi_*(x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_n^{\epsilon_n}) = \psi(x_1)^{\epsilon_1} \cdots \psi(x_n)^{\epsilon_n}. \tag{1.5.3}$$

This proves that the free group generated by X is alternatively defined up to an isomorphism by this *universal property*. The only subtlety with using universal constructions is that we need to be sure that the corresponding universal objects actually exist in their categories, and in many cases the easiest way to show it is to construct one explicitly (not up to iso), like we did for the free groups at first.

Finally, we note that the correspondence $X \rightarrow F_X$ is *natural*, i.e. functorial. Namely, any map between sets $f : X \rightarrow Y$ generates a unique homomorphism between the corresponding free groups $F_f : F_X \rightarrow F_Y$ defined on the generators by $F_f(i_X(x)) = i_Y(f(x))$. Sometimes mathematicians describe equations of this form as “ f intertwines the action of i_X and i_Y ”. Of course, this is a bit of an abuse of terminology, since f and F_f are not quite the same, although f can be thought of as a restriction of F_f . In any case, in the end we have a functor $F : \text{Set} \rightarrow \text{Gr}$.

This kind of universal construction is the most typical way “universal properties” or “characteristic properties” come up in mathematics – as universal objects in categories of morphisms with one fixed end. Indeed, as we have seen, concepts such as surjections and injections can be defined this way. This will also apply to geometric concepts such as submersions, lifting of paths, etc. For more examples and a general description see [28, Exercise 7.8:24 and Def. 7.8:12].

Remark 1.5.10. In group theory, especially in finite groups, groups are often described by a list of *generators* and *relations*. If $\{g_1, \dots, g_n\}$ is the set of generators, then a relation is a condition of the form $g_{i_1}^{\epsilon_1} \cdots g_{i_k}^{\epsilon_k} = e$. Then our group is understood to be the quotient of the free group $F_{\{g_i\}}$ by the smallest normal subgroup containing all left-hand sides of the relations. In other words, every section of any “word” in the free group that coincides with the left-hand side of a relation (or its inverse) can be replaced by the empty string.

For example, the cyclic group S_3 of permutations of the tuple $(1, 2, 3)$ can be described as a group generated by the elementary transpositions $\{\sigma_{12}, \sigma_{23}, \sigma_{31}\}$ with relations $\sigma_{ij}^2 = e$ and two extra relations

$$\begin{aligned}\sigma_{12}\sigma_{23}\sigma_{12} &= \sigma_{23}\sigma_{12}\sigma_{23}, \\ \sigma_{23}\sigma_{13}\sigma_{23} &= \sigma_{13}\sigma_{23}\sigma_{13},\end{aligned}\tag{1.5.4}$$

also known as *braid relations* in the theory of braid groups. The notation for this *presentation* of the group is

$$S_3 = \left\langle \sigma_{12}, \sigma_{23}, \sigma_{13} \mid \sigma_{ij}^2, \sigma_{12}\sigma_{23}\sigma_{12} = \sigma_{23}\sigma_{12}\sigma_{23}, \sigma_{23}\sigma_{13}\sigma_{23} = \sigma_{13}\sigma_{23}\sigma_{13} \right\rangle.\tag{1.5.5}$$

Without the first set of relations ($\sigma_{ij}^2 = e$), this would be the braid group B_3 .

Example 1.5.11 (Tensor algebra). Let V be a K -vector space of dimension $\dim V = n$. Pick an arbitrary basis $\{e_i\}_{i=1}^n$ of V . Its (contravariant) tensor algebra is explicitly constructed as

$$T(V) = \bigoplus_{k=1}^{\infty} V^{\otimes k},\tag{1.5.6}$$

where $V^{\otimes k}$ is the tensor power of V ,¹⁰ formally defined as the vector space of dimension kn spanned by the basis consisting of kn symbols

$$\{e_{i_1} \otimes e_{i_2} \otimes \dots \otimes e_{i_k}, i_j \in \{1, 2, \dots, n\}, j = 1, \dots, k\}.\tag{1.5.7}$$

The multiplication operation \otimes on $T(V)$ is just defined on this large basis as the concatenation of these strings of tensor products of e_i 's, and extended to all of $T(V)$ by linearity.

Note how strongly this construction depends on the initial choice of basis on V . We call such constructions *unnatural*, because they rely on a lot of extra choices, and require a proof that the result, up to isomorphism, is independent of these choices. To give a natural construction of $T(V)$, we need to discuss multilinear maps. A k -linear map $f : V^k \rightarrow A$ (where $V^k = V^{\times k} = V^{\oplus k}$) is just the direct power of V and A is a K -vector space or a

K -algebra), is a map satisfying for all $\alpha, \alpha' \in K$

$$\begin{aligned} f(v_1, \dots, v_{i-1}, \alpha v_i + \alpha' v'_i, v_{i+1}, \dots, v_k) &= \\ = \alpha f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_k) + \alpha' f(v_1, \dots, v_{i-1}, v'_i, v_{i+1}, \dots, v_k) \end{aligned} \quad (1.5.8)$$

for every i , i.e. it is K -linear in every argument separately. Note that such a map is not linear on V^k . E.g. $f(v, w) = v^1 w^2 - v^2 w^1$ is bilinear but obviously not linear as a function of $(v, w) \in \mathbb{R}^2 \oplus \mathbb{R}^2$. Multilinear functions are also known as *forms* (in particular, k -forms).

We claim that every k -linear map can be uniquely identified with a *linear* map $f_* : V^{\otimes k} \rightarrow A$. Namely, we define its action on the basis by

$$f_*(e_{i_1} \otimes \dots \otimes e_{i_k}) = f(e_{i_1}, \dots, e_{i_k}), \quad (1.5.9)$$

and then extend it to all of $V^{\otimes k}$ by linearity. Now, this definition is self-consistent, because

$$\alpha \cdot (e_{i_1} \otimes \dots \otimes e_{i_k}) = (\alpha e_{i_1}) \otimes \dots \otimes e_{i_k}, \quad (1.5.10)$$

so by k -linearity of f ,

$$f_*(\alpha \cdot e_{i_1} \otimes \dots \otimes e_{i_k}) = \alpha f_*(e_{i_1} \otimes \dots \otimes e_{i_k}), \quad (1.5.11)$$

proving that f_* respects multiplication by scalars. Therefore k -linear functions on V are the same as linear functions on $V^{\otimes k}$!

This observation can be reformulated as a *universal property* for tensor products. Let V and W be two vector spaces and let $f : V \oplus W \rightarrow U$ be a *bilinear* function on the direct sum taking values in a vector space U . Then we can define the vector space $V \otimes W$ by the universal property that any such f must factor through $V \otimes W$ via a universal linear map $t : V \oplus W \rightarrow V \otimes W$ and a unique *linear* map $f_* : V \otimes W \rightarrow U$:

$$\begin{array}{ccc} & V \oplus W & \\ t \swarrow & \curvearrowright & \searrow \forall f \text{-bilinear} \\ V \otimes W & \dashrightarrow & U \\ & \exists! f_* \text{-linear} & \end{array} \quad (1.5.12)$$

Our argument above shows that in that explicit construction, if we pick $t(v, w) = v \otimes w$ then the commutation of the triangle says $f_*(v \otimes w) = f(v, w)$, which is exactly what we had before. Thus our construction with this t indeed has this universal property and defines a unique isomorphism class of objects that are collectively denoted by $V \oplus W \xrightarrow{t} V \otimes W$. Note

that the above diagram lives completely within the category of vector spaces: all arrows are just linear maps, i.e. morphisms of vector spaces.

Now we can finally extend this construction to find the universal property of $T(V)$. As a vector space, $T(V) = \bigoplus_{k \geq 1} V^{\otimes k}$, and we know that linear functions on every $V^{\otimes k}$ need to correspond to multilinear maps on V . However, since $T(V)$ is supposed to be a tensor algebra and not just a vector space, we need to capture this extra structure in the requirements for our morphisms.

We claim that the tensor algebra $T(V)$ is defined up to unique isomorphism by the universal property that any *linear, algebra-valued* function $f : V \rightarrow A$ on V factors through $T(V)$ via a universal map $i : V \rightarrow T(V)$ and a unique *algebra homomorphism* $f_* : T(V) \rightarrow A$:

$$\begin{array}{ccc} & V & \\ i \swarrow & \curvearrowright & \searrow \forall f - \text{linear} \\ T(V) & \dashrightarrow & A \\ \exists! f_* - \text{algebra hom.} & & \end{array} \quad (1.5.13)$$

Indeed, we have already constructed this f_* on the individual summands in (1.5.6), can extend it to all of $T(V)$ by linearity, and turn it into a true algebra homomorphism by mapping tensor products in $T(V)$ to algebra products in A . It is easy to check that f_* is uniquely defined by the requirement that the triangle commute, since $i : V \rightarrow T(V)$ in our explicit example is nothing but the inclusion of V as the first direct summand in (1.5.6), and the values of f_* on V uniquely define f_* on all of $T(V)$.

Exercise 1.7. Check that the correspondence $V \mapsto T(V)$ extends to the covariant *tensor functor* $T : \text{FinVect}_K \rightarrow \text{Alg}_K$. In other words, show that every linear map of vector spaces $\psi : V \rightarrow W$ extends to an algebra homomorphism $\psi_* : T(V) \rightarrow T(W)$. Confusingly, elements of $T(V)$ are called *contravariant tensors* on V , but T is a covariant functor.

1.5.3 Adjoint functors

Definition 1.5.12 (Adjoint functors). Let C and \mathcal{D} be two categories. Let also $F : C \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow C$ be two covariant functors. F is called the *left adjoint* of G (and G the right adjoint of F), denoted by ${}^*G = F$ (and $F^* = G$ respectively) if

1. For every pair of objects $X \in \text{Ob}(C)$ and $Y \in \text{Ob}(\mathcal{D})$ we have a bijection of sets

$$\varphi_{X,Y} : \text{Hom}_{\mathcal{D}}(F(X), Y) \cong \text{Hom}_C(X, G(Y)). \quad (1.5.14)$$

2. The above bijections are *natural* in X and Y , which means that for any morphisms

$f : X' \rightarrow X$ and $g : Y \rightarrow Y'$ the following square commutes:

$$\begin{array}{ccc}
 \text{Hom}_{\mathcal{D}}(F(X), Y) & \xrightarrow{\varphi_{X,Y}} & \text{Hom}_C(X, G(Y)) \\
 \text{Hom}_{\mathcal{D}}(F(f), g) \downarrow & \curvearrowleft & \downarrow \text{Hom}_C(f, G(g)) \\
 \text{Hom}_{\mathcal{D}}(F(X'), Y') & \xrightarrow{\varphi_{X',Y'}} & \text{Hom}_C(X', G(Y'))
 \end{array} \quad (1.5.15)$$

(here we use Hom as a functor of two arguments, which we haven't defined, but it suffices to fix X and vary Y , or vice versa, so Hom becomes a regular hom-functor, see example 1.2.3).

The naturality requirement makes sure that $\varphi_{X,Y}$ respects the categorical structures of C and \mathcal{D} . Also note that $F^* \neq {}^*F$ in general. Finally, for contravariant functors there are two different types of adjoint functors, but we will not go into that here. For example, contravariant functors are often adjoint to themselves (like e.g. $\text{Hom}(_, Z)$).

Proposition 1.5.13 (Uniqueness of adjoints). *If an adjoint functor (left or right) exists, then it is unique up to natural isomorphisms.*

Proof. Suppose both G and G' are the right adjoints of F . Then there is a natural bijection $\text{Hom}(X, G(Y)) \cong \text{Hom}(X, G'(Y))$ given by $\varphi_{X,Y}$. Denote $Z = G(Y)$ and $\tau_Y = \varphi_{G(Y), Y}(\text{id}_Y)$. We claim that τ_Y are the components of a natural isomorphism between G and G' . This can be checked by taking a morphism $f : Y \rightarrow Y'$ and following two commuting squares. First trace how $\text{id}_{G(Y)}$ is acted upon by this square

$$\begin{array}{ccc}
 \text{Hom}(G(Y), G(Y)) & \longrightarrow & \text{Hom}(G(Y), G'(Y)) \\
 \downarrow & \curvearrowleft & \downarrow \\
 \text{Hom}(G(Y), G(Y')) & \longrightarrow & \text{Hom}(G(Y), G'(Y'))
 \end{array} \quad (1.5.16)$$

to get $G'(f) \circ \tau_Y = \varphi_{G(Y), Y'}(G(f))$. Secondly, trace how $\text{id}_{G(Y')}$ around this square

$$\begin{array}{ccc}
 \text{Hom}(G(Y), G(Y)) & \longrightarrow & \text{Hom}(G(Y), G'(Y)) \\
 \downarrow & \curvearrowright & \downarrow \\
 \text{Hom}(G(Y), G(Y')) & \longrightarrow & \text{Hom}(G(Y), G'(Y')) \\
 \end{array} \tag{1.5.17}$$

to get $\varphi_{G(Y), Y'}(G(f)) = \tau_{Y'} \circ G(f)$.

Combined, we have $G'(f) \circ \tau_Y = \tau_{Y'} \circ G(f)$, showing that τ is a natural transformation. Since φ is a bijection, natural in X and Y , for each Y , τ_Y is also a bijection, which implies that τ is an isomorphism. ■

Example 1.5.14. In Set , let $F : \text{Set} \rightarrow \text{Set}$ be the functor that evaluates the Cartesian product with a fixed set Y : $F(_) = _ \times Y$ and let G be the covariant hom-functor $G(_) = \text{Map}(Y, _) : \text{Set} \rightarrow \text{Set}$. It is easy to see that we have a bijection

$$\varphi_{X, Z} : \text{Map}(X \times Y, Z) \cong \text{Map}(X, \text{Map}(Y, Z)), \tag{1.5.18}$$

which maps every Z -valued function of two variables $f(x, z)$ to its *partial function*

$$\varphi_{X, Z} : ((x, y) \mapsto f(x, y)) \mapsto (x \mapsto (y \mapsto f(x, y))). \tag{1.5.19}$$

The naturality of this set of bijections is a lot easier to check by doing it as an exercise, like most abstract proofs in category theory.

Example 1.5.15. On Ab , Vect_K , $R\text{-Mod}$ and many other categories, \otimes and Hom are adjoint:

$$\text{Hom}(U \otimes V, W) \cong \text{Hom}(U, \text{Hom}(V, W)). \tag{1.5.20}$$

Example 1.5.16. The functor $F : \text{Set} \rightarrow \text{Gr}$ which constructs free groups ($X \mapsto F_X$) is adjoint to the forgetful functor that turns a group into its underlying set.

1.5.4 Universal constructions as adjoints

In this section we will see that a lot of the familiar universal constructions can be reformulated as functors adjoint to some forgetful functors.

Idea Let \square be a forgetful functor. Some examples are $\text{Gr} \rightarrow \text{Set}$, $\text{CommRing} \rightarrow \text{Set}$, $\text{Top} \rightarrow \text{Set}$ (Top is the category of topological spaces), $\text{CommRing} \rightarrow \text{Ab}$, $\text{Ring} \rightarrow \text{Lie}$. Then a *universal construction* is a left adjoint to \square .

Example 1.5.17. $\square : \text{Gr} \rightarrow \text{Set}$. Then the free group functor $F = {}^*\square$ is characterized by

$$\text{Hom}(F(X), G) \cong \text{Map}(X, \square(G)). \quad (1.5.21)$$

This equality says that every G -valued function on X (right-hand side) uniquely extends to a group homomorphism $F(X) \rightarrow G$. This is literally the universal property of the free group F_X ! And, as we saw earlier, F is indeed a functor.

Example 1.5.18. $\square : \text{AssRing} \rightarrow \text{LieRing}$, $R \mapsto R^{(-)}$. The left adjoint of this forgetful functor is denoted by $\mathcal{U}(L)$, $L \in \text{Ob}(\text{LieRing})$, and is called the *universal enveloping ring (algebra)* of L . The universal property of the enveloping algebra states that every homomorphism of Lie rings $L \rightarrow R^{(-)}$ extends to a homomorphism of associative rings (algebras) $\mathcal{U}(L) \rightarrow R$.

In the theory of Lie algebras, if L is a Lie algebra consisting of some explicitly given matrices, then $\mathcal{U}(L)$ is the matrix algebra generated by all elements of L by taking all possible combinations of sums and matrix products (not just matrix commutators!).

Example 1.5.19. $\square : R\text{-Alg} \rightarrow R\text{-Mod}$, where R is a commutative ring. Then the left adjoint is denoted by T :

$$\text{Hom}_{R\text{-Alg}}(T(U), A) = \text{Hom}_{R\text{-Mod}}(U, \square(A)). \quad (1.5.22)$$

Again, this is literally the universal property of the tensor algebra of U , so $T(U)$ is the tensor algebra of the R -module U .

Having the tensor algebra $T(L)$ at hand, we can reconstruct the universal enveloping ring of L by just introducing extra relations $[x, y] = x \otimes y - y \otimes x$ for all x and y , i.e.

$$U(L) = T(L) / \langle [x, y] - x \otimes y + y \otimes x \mid x, y \in L \rangle. \quad (1.5.23)$$

1.6 Products and coproducts

In set theory, we have two set operations defined in the axioms themselves: $X \cup Y$ and $X \cap Y$. These operations are completely Boolean, i.e. use nothing but logical operators in their definitions.

On the other hand, Cartesian products $X \times Y$ are usually defined as sets of “ordered pairs” or “tuples” (x, y) . But what *are* ordered pairs? There is no such construction in the axioms of set theory! Traditionally this was solved by constructions like Kuratowski’s pair $(x, y) \equiv \{\{x\}, \{x, y\}\}$, because this formula unambiguously distinguishes x from y . However, we immediately run into syntactical problems when we consider (x, x) . A much bigger problem, however, is the fact that this construction could be replaced by myriads of other ways to represent ordered pairs (as it indeed was, historically). This suggests, once again,

that we should look for a different, *natural* construction of Cartesian products.

The same problem concerns disjoint unions $X \sqcup Y$, which are also not a Boolean operation and require replacing any elements in $X \cap Y$ by some other (completely arbitrary!) objects before taking the union.

$X \times Y$ is special in that it has two *canonical projections* onto the components of the product! We can use this to come up with a universal property for products. Similarly, the disjoint union $X \sqcup Y$ has two *canonical inclusions* of X and Y into the union.

Definition 1.6.1 (Products). In a general category C , a *product* $X \times Y$ of two objects is a *terminal* object in the category of diagrams in C that look like this:

$$\begin{array}{ccc} & \bullet & \\ & \swarrow & \searrow \\ X & & Y \end{array} \quad (1.6.1)$$

Here, only the vertices X and Y are fixed, whereas the top vertex and the two morphisms can be anything, and morphisms between such diagrams are morphisms between the top vertices such that all triangles commute. Then the universal property looks like this (this diagram is called a *Cartesian square*):

$$\begin{array}{ccccc} & & \bullet & & \\ & \swarrow & & \searrow & \\ X & & \exists! & & Y \\ & \nwarrow \pi_1 & \downarrow & \nearrow \pi_2 & \\ & X \times Y & & & \end{array} \quad (1.6.2)$$

meaning that for any top (unlabeled) part of the diagram there exists a unique middle morphism that makes both triangles commute. This property is very easy to check for Cartesian products in set theory. If the top morphisms are f and g , then the middle morphism is denoted $f \times g$ and its action in set theory is $(f \times g)(z) = (f(z), g(z))$, which is uniquely fixed by the commutation of the triangles.

Definition 1.6.2 (Coproducts). Similarly, a coproduct $X \sqcup Y$ is defined as an *initial* object in

the category of the following diagrams:

$$\begin{array}{ccc}
 & \bullet & \\
 X & \nearrow & \nwarrow \\
 & Y &
 \end{array}
 \quad
 \begin{array}{ccc}
 & \bullet & \\
 X & \nearrow & \nwarrow \\
 & \exists! & \\
 i_1 & \searrow & \swarrow i_2 \\
 & X \sqcup Y &
 \end{array} \tag{1.6.3}$$

If the top morphisms are f and g , then the middle morphism is denoted $f \sqcup g$ (in set theory, “piecewise function”).

Definition 1.6.3 (Pullbacks). We can further generalize the above definitions and consider the category of commuting diagrams of the shape:

$$\begin{array}{ccc}
 \bullet & \longrightarrow & Y \\
 \downarrow & \curvearrowleft & \downarrow g \\
 X & \xrightarrow{f} & Z
 \end{array} \tag{1.6.4}$$

Here, only X, Y, Z and f, g are fixed. Morphisms between such diagrams are defined as usual. The terminal object in this category is called the *pullback*, or the *fibered product*, $X \times_Z Y$. This universal property is described by the commuting diagram:

$$\begin{array}{ccccc}
 & \bullet & & & \\
 & \dashrightarrow & & & \\
 & \exists! & & & \\
 & \searrow & \nearrow & & \\
 & X \times_Z Y & \xrightarrow{\pi_2} & Y & \\
 & \downarrow \pi_1 & & \downarrow g & \\
 & X & \xrightarrow{f} & Z &
 \end{array} \tag{1.6.5}$$

meaning that for every top (unlabeled solid) part of the diagram, there exists a unique dashed morphism that makes this diagram commute (i.e. the top two triangles). Note that the pullback strongly depends on the morphisms f and g !

Definition 1.6.4 (Pushouts). The dual concept to pullbacks is the initial object, called the *pushout*, or *fibered coproduct* $X \sqcup_Z Y$, in the category of the following diagrams:

$$(1.6.6)$$

In concrete categories, where i_1, i_2 are usually *inclusion* maps, we say that there is a unique morphism that *restricts* to i_1 and i_2 on X and Y .

Example 1.6.5. In Set , the pullback is simply the *equalizer* of f and g ,

$$X \times_Z Y = \{(x, y) \in X \times Y \mid f(x) = g(y)\}, \quad (1.6.7)$$

with projections π_1 and π_2 being just the restrictions of the canonical projections in the Cartesian product.

Pushouts are more sophisticated: in set theory, it is the *coequalizer*

$$X \sqcup_Z Y = (X \sqcup Y) /_{\sim}, \quad (1.6.8)$$

where \sim is the finest (i.e. has smallest truth domain) equivalence relation under which $f(z) \sim g(z)$.

Example 1.6.6. Category Top_* of pointed topological spaces has a coproduct equal to the wedge sum $X \sqcup Y = X \vee Y$, which is the disjoint union with only the two basepoints identified (glued). The product \times in this category, on the other hand, takes the Cartesian product and identifies all points in the set $(X \times \{y\}) \cup (\{x\} \times Y)$ to make a new basepoint out of it.

Example 1.6.7. In Gr , the product is the direct product of groups, while the coproduct is the *free product* of groups

$$G \sqcup H = G * H, \quad (1.6.9)$$

which is the quotient of the free group formed by all elements of both G and H by all relations within G and within H separately. In other words, it is the “most free” group containing G and H as subgroups.

Fibered coproducts \sqcup_Z in group theory also exist and are called *amalgamated free products*. Namely $G *_K H$ is the factor of the free product by all relations of the form $f(k)g(k)^{-1} = e$. E.g. $\mathbb{Z}_4 *_{\mathbb{Z}_2} \mathbb{Z}_6 = \text{SL}(2, \mathbb{Z})$.

Example 1.6.8. In CommRing , Vect_K , Alg_K , $R\text{-Mod}$ and many others, we have the direct product $R \times S$, and the tensor product $R \otimes S$ for the coproduct.

Example 1.6.9. In Vect , Ab , $R\text{-Mod}$ (with commutative R), and more generally in any *abelian category*, products and coproducts are isomorphic and are the direct sum $\times \cong \sqcup \cong \oplus$.

Example 1.6.10. Comp is the category of compact Hausdorff topological spaces. The classical Tikhonoff theorem states that products exist not only for any two objects in this category, but for any *infinite* set of objects (of any cardinality)!

Remark 1.6.11. Geometry is hard because products are difficult to construct (see how restrictive Tikhonoff theorem is), especially fibered products (you are forced to extend the category). Conversely, in algebra, coproducts are usually very complicated objects. In a precise sense, these statements are actually dual to each other (see below).

1.7 Natural transformations and category equivalence

Definition 1.7.1 (Natural transformations of functors). For two functors (both, say, covariant) $F, G : C \rightarrow \mathcal{D}$, a natural transformation τ of F into G , written as $\tau : F \Rightarrow G$, is a collection of morphisms $\{\tau_X : F(X) \rightarrow G(X) \mid X \in \text{Ob}(C)\}$ in \mathcal{D} such that the images of any morphism $f : X \rightarrow Y$ in C under F and G together with τ_X and τ_Y form commuting squares:

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ \tau_X \downarrow & \curvearrowright & \downarrow \tau_Y \\ G(X) & \xrightarrow{G(f)} & G(Y) \end{array} \quad (1.7.1)$$

The higher-level depiction of natural transformations is

$$\begin{array}{ccc} & F & \\ & \swarrow \tau \downarrow & \searrow \\ C & & \mathcal{D} \\ & \curvearrowright & \\ & G & \end{array} \quad (1.7.2)$$

Remark 1.7.2. There are two ways to compose natural transformations, namely the *horizontal*

composition $\tau \oplus \rho$:

$$\begin{array}{ccc}
 F_1 & & F_2 \\
 \textcirclearrowleft \downarrow \tau \textcirclearrowright & & \textcirclearrowleft \downarrow \rho \textcirclearrowright \\
 C & \xrightarrow{\quad} & \mathcal{D} \\
 G_1 & & G_2 \\
 & \xrightarrow{\quad} & \mathcal{E}
 \end{array} \rightsquigarrow \begin{array}{c}
 F_1 F_2 \\
 \textcirclearrowleft \downarrow \tau \oplus \rho \textcirclearrowright \\
 C \\
 G_1 G_2 \\
 \mathcal{E}
 \end{array} \quad (1.7.3)$$

and the *vertical composition* $\tau \ominus \rho$:

$$\begin{array}{ccc}
 F \\
 \textcirclearrowleft \downarrow \tau \textcirclearrowright \\
 C & \xrightarrow{\quad} & \mathcal{D} \\
 G & \xrightarrow{\quad} & \\
 \textcirclearrowleft \downarrow \rho \textcirclearrowright \\
 H
 \end{array} \rightsquigarrow \begin{array}{c}
 F \\
 \textcirclearrowleft \downarrow \tau \ominus \rho \textcirclearrowright \\
 C \\
 G \\
 \mathcal{D}
 \end{array} \quad (1.7.4)$$

These two compositions are both associative on their own, and in addition to that they satisfy the “2-associativity” law

$$(\pi \ominus \rho) \oplus (\sigma \ominus \tau) = (\pi \oplus \rho) \ominus (\sigma \oplus \tau),$$

$$\bullet \textcirclearrowleft \downarrow \textcirclearrowright \bullet \textcirclearrowleft \downarrow \textcirclearrowright \bullet \quad (1.7.5)$$

which illustrates that the double composition of the above diagram should be independent of the order of compositions.

This allows us to define *2-categories*: they consist of an Ob class, 1-morphisms Mor_1 acting between objects with an associative composition, and 2-morphisms Mor_2 acting between 1-morphisms with two associative and 2-associative compositions. However, now that we have a notion of isomorphism of functors (via natural isomorphisms), there is no reason to require literal associativity of the composition of morphisms and literal equality of $1_{A,f}$ (or $f1_A$) and f . Instead, we require that there be a natural isomorphism τ (*associator*) such that $\tau_{X,Y,Z} ((fg)h) = f(gh)$ and $u_{X,Y}^{L/R}$ (*unit*) such that $u_{X,Y}^L (1_X f) = f$ and $u_{X,Y}^R (f 1_Y) = f$. For example, the category of all categories with Mor_2 being the classes of natural transformations is a 2-category.

Obviously, this can be extended to *n-categories*, where at each level we have *n* different compositions that need to be “*n*-associative”. And similarly to the above, in an *n*-category, all laws imposed on Mor_k with $k < n$ should be understood only up to $(k + 1)$ -isomorphism. *n*-Categories can be visualized as *n*-dimensional volumes: 1-morphisms are lines, 2-morphisms are membranes stretched between two lines, 3-morphisms are 3-dimensional volumes between two membranes with common edges, etc. In theoretical computer science, category theory replaces *type theory*: the main category is the category of all types of data (Int, Float, List, String are examples of objects in this category), 1-morphisms are functions/programs that convert types to types, 2-morphisms are functions that rewrite functions, etc. See [2].

Example 1.7.3. The identity transformation $1_F : F \Rightarrow F$, $\tau_X = \text{id}_{F(X)}$.

Example 1.7.4. Consider again the covariant power set functor, which we will denote this time simply by $P : \text{Set} \rightarrow \text{Set}$. Any set X can be embedded inside $P(X)$ simply as singletons $x \mapsto \{x\}$. Denote this embedding by σ_X . This map is natural in the following sense. Given a map between sets $f : X \rightarrow Y$, we have $\sigma P(f) = f\sigma$, i.e. the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \sigma_X \downarrow & \curvearrowright & \downarrow \sigma_Y \\ P(X) & \xrightarrow{P(f)} & P(Y) \end{array} \quad (1.7.6)$$

This can be interpreted as saying that the embedding σ_X looks the same regardless of any relabeling of the elements of X (is equivariant with respect to bijections on X). In formal language, this diagram represents a natural transformation from the identity functor to P : $\text{id}_{\text{Set}} \Rightarrow P$. This is what we mean when we say that the embedding σ is *natural* or *canonical* (canonical usually implies uniqueness but it's not a strictly defined term in category theory).

Example 1.7.5. $C = \mathcal{D} = \text{FinVect}_K$. We have the identity functor 1_C and the double dual functor $F : V \rightarrow V^{**}$. A *classical theorem in linear algebra* states that V^{**} is naturally isomorphic to V , i.e. there is a natural transformation $1_C \Rightarrow F$ that is invertible (i.e. every τ_V is an isomorphism and $(\tau^{-1})_V = (\tau_V)^{-1}$).

Explicitly, the isomorphism is given by $v \mapsto v^{**}$, where $v^{**} \in V^{**}$ is defined as a functional on V^* by the formula $v^{**}(\alpha) = \alpha(v)$ for any $\alpha \in V^*$. This formula does not contain any arbitrary choices, which makes it natural intuitively. Indeed, it is a natural isomorphism because the square formed by this τ_V and any automorphism of V and its double dual commutes.

Contrast this with the statement that V is isomorphic to V^* . The only way to construct an isomorphism of this kind is to explicitly introduce a basis $\{e_i\}_{i=1}^n$ on V and define its dual basis $\{\alpha^i\}_{i=1}^n$ by $\alpha^i(e_j) = \delta_j^i$ (or any other basis of V^*). Then we have an isomorphism $V \rightarrow V^*$ that maps $e_i \mapsto \alpha^i$. However, this isomorphism is *not natural* because it strongly depends on

the choice of basis, i.e. it is not invariant under isomorphisms of V . In categorical language, there simply are no natural transformations between a covariant functor (the identity functor) and a contravariant one (the dualization functor). Even if we were to define what's called *dinatural transformations* (where the analogous square with only the arrow $G(Y) \rightarrow G(X)$ reversed must commute), one can show that the only such transformation from the identity to the dualization functor is the zero transformation.

Note that the situation is much more complicated in the infinite-dimensional case. For example, the space dual to the space of compactly supported scalar functions on a locally compact Hausdorff space is the space of *measures*. Any infinite-dimensional space can always be embedded in its double dual, and sometimes shown to be strictly “smaller” (e.g. when this embedding is a proper map). However, there is no obvious embedding of the original space into its single dual in these examples at all.

Exercise 1.8. Formalize the statement that there *is* a “canonical” isomorphism between a vector space with a metric (or any other non-degenerate pairing of vectors) and its dual. In other words, non-natural constructions can become natural in a properly “enriched” category, where objects have more structure.

Exercise 1.9 (From [29]). Let FinBij be the category of finite sets and bijections. Given $X \in \text{FinBij}$, let $\text{Sym}(X)$ be the set of permutations on X (the set of all bijections $\pi : X \rightarrow X$) and let $\text{Ord}(X)$ be the set of all total orders on X (which are essentially just strings $x_0 \leq x_1 \leq \dots \leq x_n$). Extend Sym and Ord to functors $\text{FinBij} \rightarrow \text{Set}$ (i.e. provide an arrow map, an object map has already been given) *in a natural way* (it is important that FinBij is equipped with bijections only whereas Set is not). Exhibit isomorphisms $\text{Sym}(X) \simeq \text{Ord}(X)$ for every finite set X , but show that it is *impossible* for there to exist a natural isomorphism (or indeed, any natural transformation at all!) $\alpha : \text{Sym} \rightarrow \text{Ord}$. *Hint:* consider identity permutations.

Definition 1.7.6 (Skeleton). A category is called *skeletal* if it does not have any nontrivial isomorphisms, i.e. two objects are isomorphic iff they are equal. For any category C , its *skeleton* $\text{Sk}(C)$ is a *full* (i.e. all morphisms between surviving objects are left intact) skeletal subcategory of C such that every object in C is isomorphic to some object in $\text{Sk}(C)$. (Essentially the skeleton is a “minimal” subcategory that is equivalent to C . Notice that skeletons are defined only up to isomorphisms of categories.)

Example 1.7.7. $\text{Ob}(\text{Sk}(\text{Set})) = \text{Ord}$, the class of cardinal numbers. $\text{Ob}(\text{Sk}(\text{FinSet})) = \mathbb{N}$.

Example 1.7.8. $\text{Ob}(\text{Sk}(\text{Vect}_K))$ consists of all powers of K , K^α , where α is any cardinal number. $\text{Ob}(\text{Sk}(\text{FinVect}_K))$ is the set of “column-vector spaces” K^n .

Example 1.7.9. In Set , the functor $\text{Map}(_, \times _, _)$ is naturally isomorphic to $\text{Map}(_, \text{Map}(_, _))$ (order of arguments preserved).

Definition 1.7.10 (Isomorphism of categories). Categories C and \mathcal{D} are called *isomorphic* if there exists a pair of functors $F : C \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow C$ such that $FG = 1_C$ and $GF = 1_{\mathcal{D}}$. The notation is $C \cong \mathcal{D}$. They are called anti-isomorphic if $C \cong \mathcal{D}^*$.

Example 1.7.11. $\mathbb{Z}\text{-Mod} \cong \text{Ab}$. Indeed, every \mathbb{Z} -module is turned into an abelian group by the forgetful functor that forgets the multiplication map $\mathbb{Z} \times M \rightarrow M$. And vice versa, the multiplication map can be recovered from any abelian group by defining $n \cdot x = x + x + \dots + x$. By applying both functors in succession one obtains not just isomorphic, but identical objects, so this is indeed an isomorphism of categories.

However, the notion of isomorphic categories is entirely useless in practice, because in reality we want to allow throwing out extraneous isomorphic objects without considering the resulting category different.

Definition 1.7.12 (Equivalence of categories). Two categories are equivalent, $C \simeq \mathcal{D}$, if there are two functors $F : C \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow C$ such that the following are isomorphic as functors, i.e. have natural isomorphisms: $FG \cong 1_C$ and $GF \cong 1_{\mathcal{D}}$. Two categories are anti-equivalent if $C \simeq \mathcal{D}^*$.

Proposition 1.7.13. $C \simeq \mathcal{D}$ iff $\text{Sk}(C) \cong \text{Sk}(\mathcal{D})$.

Definition 1.7.14 (Natural isomorphism). A natural transformation $\tau : F \Rightarrow G$ is called a natural isomorphism (or an isomorphism of functors) if it has a two-sided inverse $\tau^{-1} : G \Rightarrow F$. Equivalently, each morphism τ_X is an isomorphism.

Example 1.7.15. Consider a special case of the quotient functor $\text{GrPair} \rightarrow \text{Gr}$ that we constructed in Example 1.3.2. Namely, restrict the category GrPair to have only morphisms $f : (G_1, H_1) \rightarrow (G_2, H_2)$ such that $\ker f = H_1$. Let us call thus category Ker . Then we have a functor that maps (G, H) to $(G/H, G, \pi)$ where π is the natural projection (quotient map) $G \rightarrow G/H$. This is a functor into the category Fact of triples (K, G, π) where π is a group epimorphism $G \rightarrow K$. Morphisms are pairs of homomorphisms on K and G that form commuting squares with the projections. Then we can invert this functor (up to natural isomorphism), namely an object (K, G, π) goes to $(G, \ker \pi)$ and a morphism $f_* : (K_1, G_1, \pi_1) \rightarrow (K_2, G_2, \pi_2)$ can be mapped to $f = \pi_1 f_* = f_* \circ \pi_1 : G_1 \rightarrow G_2$. Clearly $\ker f = \ker \pi_1$, so we end up in the category we started with. It is easy to check that a sequential application of these two functors in any order produces an isomorphic object. Thus we have restated the first isomorphism theorem as an equivalence of two categories:

$$\text{Ker} \simeq \text{Fact}. \tag{1.7.7}$$

Exercise 1.10. Extend the last example to cover the general fundamental homomorphism theorem.

Example 1.7.16. *The fundamental theorem of covering spaces*, which we will cover later (Theorem 4.4.2), is the statement that the category of covering spaces of a space X , denoted by Cov_X , is equivalent to the category of sets with an action of the fundamental group of X , denoted by $\pi_1(X)$ -Set:

$$\text{Cov}_X \simeq \pi_1(X)\text{-Set}. \quad (1.7.8)$$

Example 1.7.17. A lot of geometric categories are anti-equivalent to algebraic categories. For instance, we have *Gelfand's theory of commutative C^* -algebras*:

$$\text{CompHaus} \simeq (\text{Ban})^*, \quad (1.7.9)$$

where **CompHaus** is the category of compact Hausdorff topological spaces, and **Ban** is the category of commutative complex Banach C^* -algebras. Explicitly, a topological space X on the left corresponds to the algebra $C(X)$ of continuous complex-valued functions on the right. In fact, all basic differential geometry can be rewritten as a collection of statements about (sufficiently nice) commutative algebras (see e.g. [3]). Moreover, this principle was used to define *non-commutative geometry* as the dual of some categories of non-commutative algebras [4].

Example 1.7.18. A fundamental theorem of group theory, and the foundational theorem of *harmonic analysis*, is the *Pontryagin duality*. Its generalization to non-abelian groups (known as Tannaka-Krein duality) can be formulated as the anti-self-equivalence of the category of locally compact groups:

$$\text{LocCompGr} \simeq (\text{LocCompGr})^*, \quad (1.7.10)$$

where the mapping is $G \mapsto \hat{G}$, \hat{G} being the dual group. For instance, $\hat{\mathbb{S}^1} = \mathbb{Z}$, which means that the Fourier transform maps functions on the circle to functions on \mathbb{Z} (i.e. sequences of numbers). Similarly, $\hat{\mathbb{R}} = \mathbb{R}$, so Fourier transforms here are continuous. Pontryagin's theory shows how to define Fourier transforms for functions on an arbitrary topological abelian group.

See [1, II.2] for more details on these and other examples of equivalences of categories.

Theorem 1.7.19 (Yoneda Lemma). *Let C be a category and let $h^A = \text{Hom}_C(A, _) : C \rightarrow \text{Set}$ be a covariant representable (Hom-) functor. Let also $F : C \rightarrow \text{Set}$ be a covariant functor. Then all natural transformations $\tau : h^A \Rightarrow F$ are in one-to-one correspondence with the elements of the set $F(A)$, and this bijection is natural in A and F . Namely, τ corresponds to the element $\tau_A(\text{id}_A) \in F(A)$ (here we view $\text{id}_A \in h^A(A)$).*

Definition 1.7.20 (Representable functors). Following Yoneda's lemma, **Set**-valued functors that are naturally isomorphic to a hom-functor (h^A or h_A) are called *representable*.

Remark 1.7.21. We don't prove this theorem here, but the moral of Yoneda's Lemma is that we can naturally realize *elements* of the sets $F(A)$ in purely categorical terms! In other words, the structure of a category can be “probed” by set-theoretical means by considering various set-valued functors.

In his Algebraic Geometry class a few years back, Ravi Vakil explained Yoneda's lemma like this: You work at a particle accelerator. You want to understand some particle. All you can do is throw other particles at it and see what happens. If you understand how your mystery particle responds to all possible test particles at all possible test energies, then you know everything there is to know about your mystery particle. [[Source](#)]

2 General Topology

2.1 Topological spaces

2.1.1 Open and closed sets

This section will be a quick review of the basic definitions and results of point-set topology. To keep this section brief, we shall skip the proofs for most results and refer the reader to any standard textbook to topology (e.g. [21]).

Definition 2.1.1 (Topology). A topology \mathcal{T} on a set X is a collection of subsets of X with the following properties

1. $\emptyset, X \in \mathcal{T}$.
2. \mathcal{T} is closed under union of arbitrary (finite, countably or uncountably infinite) number of its elements.
3. \mathcal{T} is closed under intersection of a finite number of its elements.

The elements of the topology of a set are called *open sets*, and the double (X, \mathcal{T}) is called a *topological space*.

The intuitive meaning of a topology is that we “distinguish” two points in X only if one of them can be put in an open set not containing the other. Thus, we tend to associate the property of belonging to the same “small” open set with proximity (the smaller the open set, in terms of inclusion, the “closer” the points), even though there is no notion of distance at this point. There are variations of this interpretation of topology, see Section 2.3.

Example 2.1.2. Some elementary examples of topologies:

1. Trivial topology $\mathcal{T} = \{\emptyset, X\}$. Can’t distinguish any two points.
2. Discrete topology $\mathcal{T} = 2^X$. The space “falls apart” into separate points, because even one-point sets are open.
3. Smallest non-discrete and non-trivial topology on $\{x, y\}$: $\mathcal{T} = \{\emptyset, \{x\}, \{x, y\}\}$. x can be separated from y , but y can’t be separated from x . This phenomenon is called a non- T_1 topology.
4. The standard topology on \mathbb{R} has arbitrary finite or countable disjoint unions of open intervals (a, b) for open sets. Exercise: check that this is a topology.

Proposition 2.1.3 (Open iff contains open neighborhoods). *Let (X, \mathcal{T}) be a topological space. A set $M \subset X$ is open iff for every $x \in M$ there exists an open set U_x such that $x \in U_x$ (i.e. an open neighborhood of x) and $U_x \subset M$.*

Proof. Exercise. ■

Definition 2.1.4 (Interior). The interior of a subset M of a topological space X is defined as the union of all the open sets in X that are contained in M , and denoted by $\text{int } M$.

Since the arbitrary union of open sets is open, the interior of a subset is an open set.

Definition 2.1.5 (Subspace Topology). For $M \subset X$, (X, \mathcal{T}) being a topological space, one can define the *subspace* (or *induced*) topology \mathcal{T}_M on M as the set of those subsets of M that can be written as an intersection of M with an open set, i.e.,

$$\mathcal{T}_M = \{M \cap U \mid U \in \mathcal{T}\}. \quad (2.1.1)$$

Definition 2.1.6 (Closed Set). A set $A \subset X$ is said to be closed in X if $X \setminus A$ is open.

One can equivalently define a topology on a set as the collection of closed sets instead of open sets, by using the following theorem.

Theorem 2.1.7. For X , a topological space, the following properties are true:

1. \emptyset and X are closed.
2. Arbitrary intersection of closed sets is closed.
3. Finite union of closed sets is closed.

Proof. Exercise. ■

Note that the empty set and the topological space are both open and closed.

Definition 2.1.8 (Closure of a Subset). The closure \bar{M} of a subset $M \subset X$ is the intersection of all closed subsets of X that contain M .

Exercise 2.1. 1. Show that $\bar{\mathbb{Q}} = \mathbb{R}$ in the standard topology.

2. Show that the set $\{e^{2\pi i \alpha n}\}_{n \in \mathbb{Z}}$ is dense in the unit circle $\mathbb{S}^1 \subset \mathbb{C}$ if and only if α is irrational.

Since the arbitrary intersection of closed subsets is closed, the closure of a subset is a closed set. We also have

$$\text{int } M \subset M \subset \bar{M}. \quad (2.1.2)$$

Definition 2.1.9 (Dense Subsets). A subset $M \subset X$ is said to be dense in X if $\bar{M} = X$.

Definition 2.1.10 (Separability). A topological space X is called separable if it has a countable, dense subset.

Definition 2.1.11 (Boundary). The boundary ∂M of a subset $M \subset X$ is equivalently defined by either of the following:

1. $\partial M = \bar{M} \setminus \text{int } M$, or
2. $\partial M = \bar{M} \cap \overline{X \setminus M}$.

Definition 2.1.12 (Continuous Maps). A map $f : X \rightarrow Y$ is called *continuous at a point* $p \in X$ if for every open set $U \in Y$ such that U contains $f(p)$, $f^{-1}(U)$ is open in X . f is said to be *continuous on a subset* $A \subset X$ if it is continuous on every point in A . Finally, f is simply called *continuous* if the pre-image of any open subset of Y is open in X (equivalent to being continuous at every point of X).

Note that replacing ‘open’ by ‘closed’ in the above definition gives an equivalent definition for continuous maps. The set of continuous maps from X to Y is denoted as $C(X, Y)$. We can now define the category of all topological spaces \mathbf{Top} given by

1. $\text{Ob}(\mathbf{Top}) = \text{All topological spaces.}$
2. $\text{Mor}(X, Y) = C(X, Y)$, the set of continuous maps between the two spaces.

Definition 2.1.13 (Homeomorphism). A continuous map is defined to be a homeomorphism if it has a continuous inverse.

Homeomorphisms are isomorphisms in the category of topological spaces. Properties of topological spaces that are preserved by homeomorphisms (and more generally, by homotopies, which are essentially continuous deformations as opposed to isomorphisms, see Section 2.4) are called *topological invariants*. Topological invariants can take values in words (“connected”/“disconnected”, “compact”/“non-compact”), in integers (winding numbers, Chern numbers, Euler characteristic, genus, etc.), in groups (homotopy groups, homology groups, etc.), and even in more complicated classes like algebras, bundles, differential forms... The only way to show that two spaces are *not* homeomorphic is to find a topological invariant that is different between the two spaces.

The problem topology serves to solve is:
Are two given topological spaces homeomorphic/homotopic or not?
If not, what *topological invariant* distinguishes them?

(2.1.3)

Since there is no universal algorithm that allows one to answer the first question, by far most of the work in topology goes into constructing new topological invariants that help us define certain classes of topological spaces. The part of topology that is especially efficient in finding these invariants is *algebraic topology*, which we will not discuss in this part (but is the subject of Part IV).

Remark 2.1.14. In some applications, like knot theory, we are interested in *isotopies* of embedded topological subspaces, which are essentially smooth deformations of embeddings without self-intersections. Isotopy is an even stronger condition than homotopy or homeomorphism, which allows for more interesting invariants like knot polynomials, which would not be invariant under homotopies or homeomorphisms (every knot can be “untied” by a homeomorphism!). In particular, there are more chances to find a complete system of isotopy

invariants for certain classes of embeddings, therefore completely classifying all such objects by a fixed set of topological invariants.

Definition 2.1.15 (Comparison of topologies). Let τ, ρ be two topologies on a set X . If the inclusion

$$\tau \subset \rho \tag{2.1.4}$$

holds, we say that ρ is a finer topology, and τ is coarser. The trivial topology is the coarsest and the discrete topology is the finest. We have the inclusions $C((X, \tau), Y) \subset C((X, \rho), Y)$ and $C(Y, (X, \rho)) \subset C(Y, (X, \tau))$.

Definition 2.1.16 (Initial/Final topology). Let X, Y be two sets and $f : X \rightarrow Y$ a map between them. If X has a topology then the finest topology on Y under which f is still continuous is called the *final* topology. Similarly, if Y has a topology, then the coarsest topology on X under which f is continuous is called the initial topology. Similarly, a family of functions $f_i : X_i \rightarrow Y$ defines a final topology on Y , and a family $f_i : X \rightarrow Y_i$ defines an initial one on X .

For a set X and an equivalence relation \sim on X , one can define the quotient set X/\sim as the set of equivalence classes under the relation, i.e.,

$$X/\sim = \{[x] \mid x \in X\}, \tag{2.1.5}$$

where $[x]$ is the equivalence class of the element x defined as

$$[x] = \{u \in X \mid u \sim x\}. \tag{2.1.6}$$

We also have the canonical map $\pi : X \rightarrow X/\sim$ that takes an element x to its equivalence class $[x]$. If X is equipped with a topology, the quotient can always be equipped with an induced topology as follows.¹¹

Definition 2.1.17 (Quotient/Factor Topology). The quotient topology on X/\sim is the collection of sets in the quotient space whose inverse image under the canonical projection π is open in X , i.e.,

$$\mathcal{T}_{X/\sim} = \{U \in X/\sim \mid \pi^{-1}(U) \in \mathcal{T}_X\}. \tag{2.1.7}$$

In other words, it is the finest topology on X/\sim that makes π continuous.

The quotient space together with the canonical projection can be instead defined by its *universal property*: any continuous function $f : X \rightarrow Z$ (valued in any topological space Z)

¹¹Existence of quotients is a pretty strong property for a category. In group theory, as we know, quotients exist only with respect to normal subgroups (i.e. $x \sim y$ iff xy^{-1} is contained in a given normal subgroup), whereas in topological groups quotients don't even exist for all normal subgroups. In particular, the first isomorphism theorem doesn't hold for topological groups.

that is constant on the equivalence classes in X uniquely factors through π , so $f = f_* \circ \pi$, where f_* is the unique continuous map satisfying this condition (we say that “ f descends to X/\sim ” as f_*). This universal property also implies that the quotient topology on X/\sim is unique.

Exercise 2.2. Check that the topology constructed above for X/\sim indeed satisfies the universal property (i.e. f_* is continuous).

The quotient topology can also be defined in an equivalent way by first defining quotient maps.

Definition 2.1.18 (Quotient Map). A surjective map $\pi : X \rightarrow Y$ is called a *quotient map* if a subset U of Y is open in Y iff $\pi^{-1}(U)$ is open in X . In other words, a quotient map is a surjective, strongly continuous map. Another way to put this is that the topology on Y must coincide with the final topology defined by π .

Now, for any surjective map π from a topological space X to a set Z , we can show that there exists exactly one topology on Z for which π is a quotient map. This topology is called the quotient topology on Z . To connect with our previous definition of the quotient topology, define $x_1 \sim x_2$ iff $\pi(x_1) = \pi(x_2)$. The quotient topologies defined either way coincide with each other.

Example 2.1.19. 1. Prove the statement in the last paragraph.

2. Consider the closed unit ball $\overline{B^n} \subset \mathbb{R}^n$ with the subset topology induced from \mathbb{R}^n and define an equivalence which identifies all the boundary points $\partial B^n = \mathbb{S}^{n-1}$ into a single element, without grouping any of the interior points. The quotient space defined by this partition can be shown to be homeomorphic to the sphere \mathbb{S}^n . Thus $B^n/\sim \cong \mathbb{S}^n$.
3. The Möbius band can be defined as the factor of the standard square $[0, 1] \times [0, 1]$ by the equivalence $(0, y) \sim (1, 1 - y)$ (and any point with $0 < x < 1$ is equivalent only to itself).

From the above example, one can see that quotienting a topological space is equivalent to *gluing together* equivalent points of the topological space.

2.1.2 Generated topologies

For most topological spaces, it is possible to specify the topology using a smaller collection of open sets than the entire topology. Such a collection is called the *basis* for the topology.

Definition 2.1.20 (Basis). A collection $\mathfrak{B} \subset \mathcal{T}$ of open sets is called a basis of the topology \mathcal{T} if every open set in \mathcal{T} can be written as a union of elements of the basis: $U = \bigcup_{\alpha} B_\alpha$, $B_\alpha \in \mathfrak{B}$.

Proposition 2.1.21 (Criterion of a basis). *A collection of sets $\mathfrak{B} \subset \mathcal{T}$ is a basis for \mathcal{T} if and only if the following conditions hold:*

1. *For every $x \in X$, there is at least one basis element that contains x , i.e., \mathfrak{B} covers X .*
2. *For $x \in B_1 \cap B_2$, where $B_1, B_2 \in \mathfrak{B}$, there exists $B_3 \in \mathfrak{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.*

Proof. Exercise. ■

Proposition 2.1.22. *Given a basis of topology, one can characterize the topology as follows: a subset U of X is open iff for every point $x \in U$, there is a basis element $B \in \mathfrak{B}$ such that $x \in B \subset U$.*

Therefore in many cases it is convenient to describe topologies by only giving a basis for them.

Proposition 2.1.23. *If there exists a countable basis for a topological space (also known as second countability), the space is separable.*

Proof. Exercise. ■

Example 2.1.24. 1. One-point sets form a basis of any discrete topology.

2. The standard topology on \mathbb{R}^n can be defined by the countable basis consisting of all open balls $B_r(x)$ with rational radii $r \in \mathbb{Q}$ and centered at points with rational coordinates, $x \in \mathbb{Q}^n$.

Using the definition of a basis, we can define the topology on the Cartesian product of two topological spaces.

Definition 2.1.25 (Product Topology). For two topological spaces X and Y , the product topology on $X \times Y$ is defined as the topology generated by the basis \mathfrak{B} consisting of all sets of the form $U \times V$, where U and V are open in X and Y respectively.

Exercise 2.3. Show that the topology on \mathbb{R}^n defined in Example 2.1.24 coincides with the product topology on $\mathbb{R} \times \dots \times \mathbb{R}$, where the topology on \mathbb{R} is the standard one (generated by open intervals).

Exercise 2.4. Recall that $X \times Y$ in reality is not a set theoretical, but a categorical construction. Check the universal property for products and thus verify that this defines a product on the category **Top**.

Remark 2.1.26. The set of all pairwise Cartesian products of basis elements forms a basis for the product topology.

Definition 2.1.27 (Wedge sum). The wedge sum of two *pointed* topological spaces is defined as $X \vee Y = (X \sqcup Y)/(x_0 \sim y_0)$.

Definition 2.1.28 (Smash product). The smash product of two *pointed* topological spaces is defined as $X \wedge Y = (X \times Y)/(X \vee Y)$, where $X \vee Y$ is identified with the union of $\{x_0\} \times Y$ and $X \times \{y_0\}$.

Exercise 2.5. 1. Show that $\mathbb{S}^1 \wedge \mathbb{S}^1 = \mathbb{T}^2 / \cong \mathbb{S}^2$.
2. More generally, show that $\mathbb{S}^m \wedge \mathbb{S}^n \cong \mathbb{S}^{m+n}$.

2.1.3 Connectedness

Definition 2.1.29 (Connected Space). A space X is said to be connected if there do not exist two non-empty open sets U, V such that $U \cap V = \emptyset$ and $U \cup V = X$. Equivalently, X is connected iff the only clopen¹² subsets in it are \emptyset and X .

Definition 2.1.30 (Connected subspace). A subset of a topological space X is said to be connected if it forms a connected topological space under the subset topology.

Exercise 2.6. Show that the interval $[a, b] \subset \mathbb{R}$ is connected in the standard topology. Hint: prove that any nonempty clopen subset of the interval has to contain the whole interval.

Proposition 2.1.31 (Connectedness is a Topological Invariant). *The image of a connected space under a continuous map is connected, i.e., if $f \in C(X, Y)$ and X is connected, so is $f(X)$.*

Definition 2.1.32 (Connected Components). We define an equivalence relation on elements of a space X by declaring that $x \sim y$ if there exists a connected subspace of X that contains x and y . The equivalence classes generated by this relation are called the connected components of X .

Exercise 2.7. Show that every connected component is closed but not necessarily open.

Exercise 2.8. Show that every clopen set is a union of (possibly infinitely many) connected components.

Exercise 2.9. Show that the connected components of \mathbb{Q} (with subspace topology inherited from \mathbb{R}) are points.

The equivalence classes of the connectedness relation are also the maximal connected subsets of X . Therefore any topological space X is a disjoint union of connected components, each of which is closed.

Definition 2.1.33 (Path). A (continuous) path in a topological space X is defined as a continuous map $f : [0, 1] \rightarrow X$. A path is said to connect the points $x_0 = f(0)$ and $x_1 = f(1)$.

Definition 2.1.34 (Path-Connected Space). A topological space is path-connected if every pair of points in X can be connected by a path in X .

Similar to connected components, we can define an equivalence relation on X by stating that two points are equivalent if there exists a path connecting them. The equivalence classes of this relation are called the *path components* of X .

Proposition 2.1.35 (path-connectedness is a topological invariant). *The following properties hold for path-connected spaces:*

1. *If X is a path-connected space and $f \in C(X, Y)$, then $f(X)$ is also path-connected.*
2. *Every path-connected space is also connected.*

Proof. Exercise. ■

Note that the converse of the second point above is not true. To see why, consider the example of the *topologists' sine curve*, which is the set of points

$$X = \{y = \sin(1/x) \mid x \in (0, 1]\} \cup \{x = 0, y \in [-1, 1]\}, \quad (2.1.8)$$

as a subset of the real plane (Fig. 2.1). The set is connected, since every point on the y-axis is

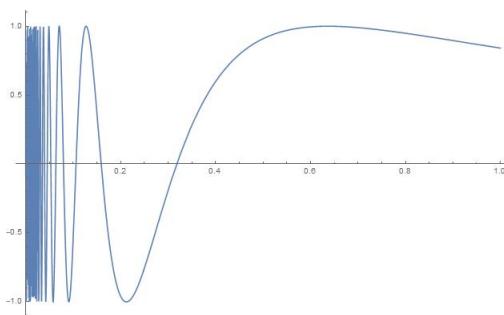


Fig. 2.1: The topologists' sine curve

a limit point of the curve. However, there is no continuous path connecting the y-axis to the rest of the curve, because $\sin(1/x)$ itself is not continuous.

- Exercise 2.10.**
1. Show that the n -dimensional sphere \mathbb{S}^n (defined as the unit sphere in \mathbb{R}^{n+1} with the subset topology) is path-connected.
 2. Show that $\mathbb{Q} \subset \mathbb{R}$ is not connected.
 3. Show that the product of two path-connected spaces is path-connected.

2.1.4 Compactness

Definition 2.1.36 (Covering/cover). A collection of subsets of a topological space X is said to cover X if the union of these subsets equals X . If the subsets are all open (or closed), the covering (also often called a cover) is said to be open (resp. closed).

Definition 2.1.37 (Compact Space). A space X is called compact if every open covering of X contains a finite subcollection that also covers X (*subcovering*).

Definition 2.1.38 (Compact Subset). A subset $A \subset X$ is called compact if it is compact in the subspace topology.

Proposition 2.1.39. *If $Y_i \subset X$, $i = 1, \dots, m$ are compact subsets, then their union $\bigcup_{i=1}^m Y_i$ is also compact.*

Proof. Exercise. ■

Proposition 2.1.40 (Compactness is a topological invariant). *The image $f(X)$ of a compact space X under a continuous map $f \in C(X, Y)$ is compact.*

Proof. Exercise. ■

- Exercise 2.11.**
1. Show that the product of two connected spaces is connected. *Hint:* first show that the union of any number of pairwise overlapping connected subspaces (i.e. subsets endowed with the subspace topology) of a space is also connected; then, write $X \times Y$ as the union of “crosses” $(\{x\} \times Y) \cup (X \times \{y\})$ that pairwise overlap and are connected only if X and Y are.
 2. Show that the product of two path-connected spaces is path-connected.
 3. Show that the product of two compact spaces is compact. *Hint:* [5].

Definition 2.1.41 (Classes of continuous maps). A map $f : X \rightarrow Y$ (not necessarily continuous) is called

1. *open* if the image of every open subset of X is open in Y ;
2. *closed* if the image of every closed subset of X is closed in Y ;
3. *proper* if it is continuous and the pre-image of any compact subset of Y is compact in X ;

- Exercise 2.12.** Show that the image of any proper map $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is closed.

2.2 Metric spaces

Definition 2.2.1 (Metric spaces). A *metric space* is a set X with a *metric* $\rho : X \times X \rightarrow \mathbb{R}$ that obeys the following conditions for any $x, y, z \in X$

1. *Non-negativity*: $\rho(x, y) \geq 0$; *non-degeneracy*: $\rho(x, y) = 0$ iff $x = y$.
2. *Symmetry*: $\rho(x, y) = \rho(y, x)$.
3. *Triangle Inequality*: $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$.

Definition 2.2.2 (Metric maps). A *metric map* between two metric spaces (X, ρ_X) and (Y, ρ_Y) is a continuous map $f \in C(X, Y)$ such that $\rho_Y(f(x_1), f(x_2)) \leq \rho_X(x_1, x_2)$. The category Met is defined as the category of metric spaces with metric maps for morphisms.

Definition 2.2.3 (Open Ball). An *open ball* $B_r(x)$ of radius r around a point x in a metric space (X, ρ) is defined as

$$B_r(x) = \{y \in X \mid \rho(y, x) < r\} \quad (2.2.1)$$

Definition 2.2.4 (Metric Topology). The metric topology induced by a metric on a set is the topology generated by basis elements that are open balls $B_r(x)$ for all $x \in X$ and $r > 0$.

Remark 2.2.5. Equivalently, by Proposition 2.1.22, a set M is open in the metric topology iff for any $x \in M$ there exists a radius $r > 0$ such that $B_r(x) \subset M$. In the metric topology, $\rho : X \times X \rightarrow \mathbb{R}$ is automatically continuous.

Proposition 2.2.6. If a metric space (X, ρ) is separable, it has a countable base.

Proof. Exercise. *Hint:* consider balls of rational radius about every point of the dense subset. ■

Definition 2.2.7 (Limit of a Sequence). A sequence of points $\{x_n\}$ in a metric space X is said to *converge* if there exists a point $x \in X$ such that for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\forall n > N, \rho(x_n, x) < \epsilon. \quad (2.2.2)$$

The point x is called the limit of the sequence.

Note that the limit of a sequence is unique by the first axiom of the metric. One can also define the limit point of a subset of a general topological space as follows:

Definition 2.2.8 (Limit Point). If A is a subset of a topological space, a point $x \in X$ is said to be a limit point of A if every open neighbourhood of x intersects A in some point other than x .

In metric spaces, limit points are equivalent to limits of sequences, i.e., for every limit point x of a subset A , there exists a sequence of points in A that converges to x . Conversely, for any convergent sequence of points in A , the limit of the sequence is a limit point of A .

Proposition 2.2.9. *For a metric space X equipped with the metric topology, the closure \bar{A} of a subset $A \subset X$ is equal to the set of limit points of A .*

Proof. Exercise. ■

Proposition 2.2.10. *For two metric spaces (X, ρ) and (Y, ρ') and a map $f : X \rightarrow Y$ such that $f(x) = y$, the following are equivalent*

1. f is continuous at x .
2. For any convergent sequence $x_n \rightarrow x$, the image is also convergent with $f(x_n) \rightarrow y$.
3. For every $\epsilon > 0$, there exists $\delta > 0$ such that $\rho(x, z) < \delta \implies \rho'(f(x), f(z)) < \epsilon$.

Proof. Exercise. ■

Definition 2.2.11 (Isometry). An isometry between two metric spaces is a bijective map $f : X \rightarrow Y$ that preserves distances, i.e.,

$$\rho_X(x, y) = \rho_Y(f(x), f(y)). \quad (2.2.3)$$

Proposition 2.2.12. *If $f : X \rightarrow Y$ is an isometry, it is also a homeomorphism.*

Proof. Exercise. ■

Definition 2.2.13 (Equivalent metrics). Two metrics ρ_1, ρ_2 on a topological space X are called equivalent if for any $x \in X$ and any $r > 0$, there exist radii $r', r'' > 0$ such that $B_{r'}^{(1)}(x) \subset B_r^{(2)}(x)$ and $B_{r''}^{(2)}(x) \subset B_r^{(1)}(x)$ (here, $B^{(i)}$ denotes an open ball defined by the metric ρ_i).

Theorem 2.2.14. *Two metrics on the same space are equivalent iff they induce the same topology.*

Proof. Exercise. ■

Definition 2.2.15 (Cauchy sequence). A sequence (x_n) of points of a metric space (X, ρ) is called a Cauchy sequence if for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\forall n, m \geq N$, $\rho(x_n, x_m) < \epsilon$.

Proposition 2.2.16. A metric space is compact iff every sequence in it has a convergent subsequence (i.e. it is sequentially compact).

Definition 2.2.17 (Completeness). A metric space (X, ρ) is called complete if every Cauchy sequence (x_n) in it has a limit $x_\infty \in X$.

Theorem 2.2.18 (Cantor's intersection theorem). Let (X, ρ) be a complete metric space and let C_n be a sequence of closed nested subsets $C_{n+1} \subset C_n$ whose diameters tend to zero, $\text{diam } C_n = \sup_{x,y \in C_n} \rho(x, y) \rightarrow 0$. Then the intersection $\bigcap_{n=1}^{\infty} C_n$ consists of exactly one point.

Proposition 2.2.19. A metric space is sequentially compact iff every decreasing sequence of closed nonempty subsets (not necessarily tending to zero diameter) has a nonempty intersection.

Theorem 2.2.20. A metric space is compact iff it is complete and for every $\epsilon > 0$ it can be covered by a finite number of open balls of radius ϵ (i.e. it is totally bounded).

Exercise 2.13. Consider \mathbb{N} with the metric $\rho(n, m) = |\frac{1}{n} - \frac{1}{m}|$. Construct a sequence of balls $B_{r_k}(x_k)$ in this space such that $r_{k+1} < r_k$, $B_{r_{k+1}}(x_{k+1}) \subset B_{r_k}(x_k)$, but the intersection of all these balls is empty.

Definition 2.2.21 (Normed vector space). A normed vector space is a vector space V over a subfield of \mathbb{C} with a function (called the norm) $\|\cdot\| : V \rightarrow \mathbb{R}_+$ such that for all $x, y \in V$:

1. $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$;
2. $\|\alpha x\| = |\alpha| \|x\|$ for any scalar (element of the structure field) α ;
3. $\|x + y\| \leq \|x\| + \|y\|$.

Every normed vector space has a naturally induced metric $\rho(x, y) = \|x - y\|$. In the corresponding induced topology, the norm is automatically continuous.

Definition 2.2.22 (Equivalent norms). Two norms $\|\cdot\|_1, \|\cdot\|_2$ on a vector space X are called equivalent if there are (universal) constants $\alpha, \beta > 0$ such that for all $x \in X$ we have $\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$.

Notice how this condition is much stronger than just the equivalence of the induced metrics, since α and β are required to be universal constants for all x .

Theorem 2.2.23. Two norms on a vector space X are equivalent iff they induce the same topology.

Proof. Exercise. ■

2.3 Separation axioms

We saw that every metric space has a canonical topology associated with it. A natural question to ask is whether the converse is true. Given a topological space, when can we define a metric on it such that the topology coincides with the metric topology? Clearly, a metric space has more structure than a topological space. In order to add more structure to topological spaces, we define the following axioms

Definition 2.3.1 (Separation Axioms). For a topological space X with arbitrary points x and y and arbitrary disjoint closed sets F and G , we define the following axioms:

1. T_1 : There exists an open neighbourhood U of x that doesn't include y , and an open neighbourhood V of y that doesn't include x , i.e., x and y are *separated*.
2. T_2 (*Hausdorff*): There exist disjoint open neighbourhoods U of x and V of y , i.e., x and y are *separated by neighbourhoods*.
3. T_3 : There exist disjoint open neighbourhoods U of x and V of F .
4. *Regular*: X is both T_1 and T_3 .
5. T_4 : There exist disjoint open neighbourhoods U of F and V of G .
6. *Normal*: X is both T_1 and T_4 .

Proposition 2.3.2. *The following results demonstrate how each axiom adds additional structure to the space:*

1. $T_1 \implies$ All finite sets are closed;
2. $T_2 \implies T_1$;
3. *Regular* \implies *Hausdorff*;
4. *Normal* \implies *Regular*.

Example 2.3.3. The category **Hausd** of Hausdorff spaces is a full subcategory of **Top**.

Theorem 2.3.4 (Urysohn Lemma). *A space X is normal iff for any two disjoint closed subsets A and B there exists a continuous function*

$$f \in C(X, [0, 1]) \text{ such that } f|_A = 0, f|_B = 1. \quad (2.3.1)$$

Theorem 2.3.5. *All metric spaces are normal.*

Proof. Exercise. ■

Theorem 2.3.6 (Urysohn). *A normal space with a countable base is metrizable, i.e., we can construct a canonical metric whose induced topology coincides with the topology of the space.*

Proposition 2.3.7. *If X is Hausdorff, every compact subset of X is closed.*

Proof. Exercise. *Hint:* prove that the complement of the compact set is open by using Proposition 2.1.3. ■

Proposition 2.3.8. *Every compact Hausdorff space is normal.*

Proof. Exercise. ■

Theorem 2.3.9 (Heyne-Borel). *In Euclidean space \mathbb{R}^n , a subset is compact iff it is closed and bounded.*

The following theorem is one of the most fundamental theorems in functional analysis, and it essentially characterizes compact subsets of the set of continuous functions defined on a compact subset of \mathbb{R}^n . We state it simply as an example of usage of the above topological terminology.

Theorem 2.3.10 (Arzelà-Ascoli). *Let X be a compact Hausdorff space. A subset F of the space $C(X, \mathbb{R})$ of continuous real-valued functions on X is relatively compact (i.e. its closure is compact) in the topology given by the sup-norm $\|f\| = \sup_X |f|$ iff it is equicontinuous (i.e. all elements $f \in F$ are uniformly continuous functions and the δ in the definition of uniform continuity is universal for all $f \in F$) and pointwise bounded (i.e. for every $x \in X$ the set $\{f(x)\}_{f \in F}$ is bounded).*

2.4 Homotopy

Definition 2.4.1 (Homotopic maps). Two continuous maps $f_0, f_1 \in C(X, Y)$ are called homotopic (we often write $f \sim g$) if there exists a continuous map, called homotopy, $F \in C([0, 1] \times X, Y)$ such that $F(0, x) = f_0(x)$ and $F(1, x) = f_1(x)$. Here, $[0, 1]$ has the standard topology. For convenience, we will often describe homotopies as *families of maps* $f_t : X \rightarrow Y$ defined by $f_t(x) = f(t, x)$.

Proposition 2.4.2. *Homotopy is an equivalence relation on $C(X, Y)$. (The set of all homotopy classes of continuous maps from X to Y is denoted by $[X, Y]$. This set defines the morphisms in the (naive) homotopy category hTop .)*

Proof. Exercise. ■

The first argument of $F(t, x)$ is essentially “time”, and $F(t, \cdot)$ continuously interpolates between f_0 and f_1 as time goes from 0 to 1. In topology, homotopy is synonymous with the phrase “continuous deformation”.

Definition 2.4.3 (Contractible space). A space X is called contractible if it is homotopy equivalent to the one-point space, i.e. there exists a point $x_0 \in X$ such that the constant map $f(x) = x_0$ is homotopic to the identity, $f \sim \text{id}_X$. This property is in fact independent of x_0 .

Proposition 2.4.4. Any contractible space is path-connected.

Proof. Exercise. Hint: consider paths traced out by $F(t, x)$ for fixed x 's. ■

Example 2.4.5. 1. $\mathbb{R}^n, B^n, (0, 1)^n$ are all contractible via, say, $F(t, x) = tx$ (they are also all homeomorphic to each other).

2. Any two continuous functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are homotopic via homotopy $F(t, x) = tg(x) + (1 - t)f(x)$.

Exercise 2.14. Show that if Y is contractible, then any two maps $f, g \in C(X, Y)$ are homotopic. This is a generalization of the above example for $X = Y = \mathbb{R}$.

Theorem 2.4.6 (Brouwer). *The sphere \mathbb{S}^n is not contractible.*

Proof. There is a plethora of different ways to prove this theorem, from pure analytical to algebraic. We will delay the proof until we compute the homology groups of the spheres in Part IV and observe that they differ from those of a contractible space (and homology groups are topological invariants). ■

Theorem 2.4.7 (Brouwer's fixed point theorem). *Any continuous map $f \in C(\overline{B^n}, \overline{B^n})$ from the closed unit ball to itself has at least one fixed point (x_0 such that $f(x_0) = x_0$).*

Proof. For $n = 1$ this is clear. For $n \geq 2$, suppose $f(x) \neq x$ for all x . Then we can define $h(x)$ as the point of intersection of the ray $[f(x), x]$ with the boundary of the ball. Then $F(t, x) = h(tx)$ is a contracting homotopy, which contradicts the last theorem. ■

Remark 2.4.8. These two theorems are in fact equivalent. The fixed point theorem can be proven purely analytically, and the contractibility theorem will follow. The simplest modern proof of non-contractibility of \mathbb{S}^n is based on homology.

Definition 2.4.9 (Retraction, Deformation retraction). A retraction of a topological space X onto its subspace $A \subset X$ is a continuous map $r : X \rightarrow A$ such that $r|_A = \text{id}_A$. A deformation retraction is a homotopy from the identity map id_X to r . A *strong* deformation retraction also

satisfies $r_t|_A = \text{id}_A$ for all $t \in [0, 1]$ (this is often included in the definition of a deformation retraction).

Definition 2.4.10 (Homotopy equivalence of spaces). Two topological spaces X and Y are called homotopy equivalent if there exist two continuous maps $f \in C(X, Y)$, $g \in C(Y, X)$ such that their compositions are homotopic to identity: $g \circ f \sim \text{id}_X$ and $f \circ g \sim \text{id}_Y$. We write $X \simeq Y$. In the presence of basepoints, these homotopies are required to leave the basepoints fixed for all values of the deformation parameter t .

Corollary 2.4.11. 1. Deformation retractions are homotopy equivalences.
2. Contractible spaces are homotopy equivalent to a point.

Example 2.4.12. 1. $\mathbb{S}^1 \times \mathbb{R} \simeq \mathbb{S}^1 \simeq M$, where M is the Möbius band.

$$2. \mathbb{R}^n \setminus 0 \simeq \mathbb{S}^{n-1}.$$

3. More generally, two spaces X, Y are homotopy equivalent iff they are both deformation retracts of the same space Z . That is, there needs to exist $A_X \subset Z$ such that $A_x \cong X$ and there exists a continuous map $F \in C([0, 1] \times Z, Z)$ such that $F_X(0, z) = z$, $F_X(1, z) \in A$ and $F_X(t, a) = a$ for any $z \in Z$, $t \in [0, 1]$ and $a \in A_x$. Similarly, there needs to exist $A_Y \subset Y$ homeomorphic to Y that is a deformation retract of Z .

2.5 The fundamental group

Definition 2.5.1 (Loops, Pointed homotopy). Let X be a path-connected topological space, pick a “base point” $x_0 \in X$, and consider the set of all *loops* in X based at x_0 , i.e. paths $\gamma \in C([0, 1], X)$ such that $\gamma(0) = \gamma(1) = x_0$. A pointed homotopy between two such loops is a homotopy $F \in C([0, 1] \times [0, 1], X)$ such that $F(\cdot, 0) = F(\cdot, 1) = x_0$. Pointed homotopy is an equivalence relation on the set of based loops.

Definition 2.5.2 (Fundamental/Poincaré group). The fundamental group $\pi_1(X, x_0)$ of a path-connected space X with a base point x_0 is defined as the set $\{[\gamma]\}$ of pointed homotopy classes of based loops with the inversion operation given by $[\gamma(t)]^{-1} = [\gamma(1 - t)]$ and the multiplication operation

$$[\gamma_1(t)] \circ [\gamma_2(t)] = \left[\begin{cases} \gamma_1(2t), & t \in [0, 1/2], \\ \gamma_2(2t - 1), & t \in (1/2, 1] \end{cases} \right]. \quad (2.5.1)$$

The unit element is given by the equivalence class of the “constant loop” $\gamma(t) \equiv x_0$. If X is not path-connected, then $\pi_1(X, x_0)$ is defined as the fundamental group of the path-connected component containing x_0 .

Exercise 2.15. Check that this definition is consistent, i.e. the homotopy classes on the right-hand sides don't depend on the choices of representatives of the classes on the left.

Definition 2.5.3 (Simple-connectedness). A path-connected space X is called simply connected if its fundamental group is trivial, $\pi_1(X, x_0) = \{e\}$.

Exercise 2.16. Show that contractible spaces are simply connected.

Theorem 2.5.4. *If a space X retracts onto a subspace A , then the homomorphism of fundamental groups $i_* : \pi_1(A, x_0) \rightarrow \pi_1(X, x_0)$ induced by the inclusion $i : A \hookrightarrow X$ is injective. If A is a deformation retract of X , then i_* is an isomorphism.*

Proof. If $r : X \rightarrow A$ is a retraction then $r \circ i = \text{id}_A$, hence by functoriality of π_1 , $r_* \circ i_* = \text{id}_{\pi_1(A, x_0)}$, which implies injectivity. If r_t is a deformation retraction, then for any loop γ in X based at $x_0 \in A$ the composition $r_t \circ \gamma$ gives a homotopy of γ to a loop in A , therefore i_* is also surjective. ■

Definition 2.5.5 (Alternative approach). Recall that the category Top_* consists of “pointed topological spaces” with chosen basepoints and morphisms that are continuous maps mapping basepoint to basepoint. Take the circle \mathbb{S}^1 and pick a point \bullet in it. Then one can redefine based loops as morphisms between two pointed topological spaces $\gamma \in C((\mathbb{S}^1, \bullet), (X, x_0))$. Pointed homotopy in the definition of π_1 can then be replaced by just homotopy of loops as maps defined on \mathbb{S}^1 . Thus, as a set, $\pi_1(X) = [\mathbb{S}^1, X]_\bullet$ (group of homotopy classes of pointed maps).

Definition 2.5.6 (Suspension). The suspension of a topological space X , is defined as the quotient space $SX = (X \times [0, 1]) / (X \times \{0\}, X \times \{1\})$ (i.e. a cylinder over X where each of the two faces is then collapsed into a point – best visualized as a double cone over X). If X is a pointed space with basepoint x_0 , then the *reduced suspension* is defined as $\Sigma X = (X \times [0, 1]) / (X \times \{0\}, X \times \{1\}, \{x_0\} \times [0, 1])$, and the new basepoint is the image of $(x_0, 0)$ in the quotient. Equivalently, $\Sigma X = X \wedge \mathbb{S}^1 = X \times \mathbb{S}^1 / (X \vee \mathbb{S}^1)$.

Exercise 2.17. Verify that Σ extends to a functor on the category of pointed spaces. This will follow from the uniqueness of the pointed map at the bottom of the following square that makes it commute:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \pi_X \downarrow & \curvearrowright & \downarrow \pi_Y \\ \Sigma X & \xrightarrow[\exists! \Sigma(f)]{} & \Sigma Y \end{array} \quad (2.5.2)$$

where π_X and π_Y are the quotient maps that define the suspensions.

Example 2.5.7. From Exercise 2.5 we have $\Sigma \mathbb{S}^n \cong \mathbb{S}^{n+1}$. This holds even when $n = 0$ and $\mathbb{S}^0 = \{-1, 1\}$.

Definition 2.5.8 (Compact-open topology). Given two spaces X, Y , denote the set of all continuous maps between them by $Y^X = C(X, Y)$. The compact-open topology on Y^X is generated¹³ by the sets $M(K, U) = \{f \in Y^X \mid f(K) \subset U\}$ where $K \subset X$ is compact and $U \subset Y$ is open.

Proposition 2.5.9. If X, Y are Hausdorff and X is locally compact, then $C(X, Y)$ is Hausdorff.

Proof. Let $f, g \in C(X, Y)$ with an $x \in X$ such that $f(x) \neq g(x)$. Since Y is Hausdorff, there exist disjoint open neighborhoods U_1 of $f(x)$ and U_2 of $g(x)$. Then $f^{-1}(U_1) \cap g^{-1}(U_2)$ is an open neighborhood of x , because f and g are continuous. Since X is locally compact, this neighborhood contains a compact neighborhood K . Then, $M(K, U_1)$ and $M(K, U_2)$ are neighborhoods of f and g , respectively. They are disjoint because U_1 and U_2 are. ■

Proposition 2.5.10. If X is a locally compact Hausdorff space, then the evaluation map $e : Y^X \times X \rightarrow Y$, defined by $e(f, x) = f(x)$, is continuous.

Proof. Let U be an open neighborhood of $f(x)$. Since f is continuous and due to the properties of X , there is a compact neighborhood K of x such that $f(K) \subset U$. Thus $f \in M(K, U)$ and $M(K, U) \times K$ is mapped into U by the evaluation map e . Finally, $M(K, U) \times K \subset Y^X \times X$ is a neighborhood of (f, x) . This proves continuity. ■

Theorem 2.5.11. Let X be a locally compact Hausdorff space, and Y and T Hausdorff spaces. For any map $f : X \times T \rightarrow Y$, denote $f_t(x) = f(x, t)$. Then the continuity of f is equivalent to each f_t being continuous separately and the mapping $t \mapsto f_t$ also being continuous from T into Y^X .

Proof. See [19, Theorem VII.2.4]. ■

This Theorem implies that if X is locally compact, then a homotopy $X \times [0, 1] \rightarrow Y$ is the same thing as a path $[0, 1] \rightarrow Y^X$.

Corollary 2.5.12 (The Exponential Law). Let X, T be locally compact Hausdorff spaces, and Y a Hausdorff space. Then the bijection from Theorem 2.5.11 is a natural homeomorphism $Y^{X \times T} \cong (Y^X)^T$ taking f to f^* defined as $f^*(t)(x) = f(x, t) = f_t(x)$.

Proof. See [19, Theorem VII.2.5]. ■

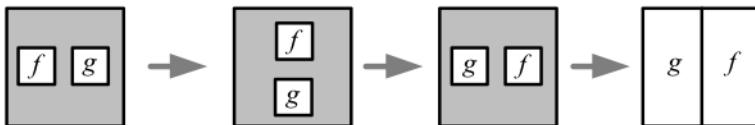
Similarly, one can show that under similar conditions, one has homeomorphisms $Y^X \times W^X \cong (Y \times W)^X$ (under $f \times g$), $Y^{X \sqcup T} \cong Y^X \times Y^T$ (under $f \sqcup g$), and $Z^Y \times Y^X \cong Z^X$ (under $f \circ g$). For details see [19, §VII.2].

Definition 2.5.13 (Loop space). The loop space of a pointed space X is the space $\Omega X = X^{\mathbb{S}^1}$ that consists of loops based at x_0 , with the compact-open topology. Its basepoint is the constant loop at x_0 .

- Proposition 2.5.14.**
1. The set of homotopy classes $[X; \Omega Y]$ is a group under the multiplication induced by the usual multiplication of loops in Y ;
 2. The set of homotopy classes $[\Sigma X; Y]$ is a group under the multiplication induced by the concatenation of maps along the interval $[0, 1] \times \{x_0\}$ in ΣX (namely $f(2t, x)$ for $t \leq 1/2$ and $g(2t - 1, x)$ after that);
 3. On the set $[\Sigma X; \Omega Y]$, the two multiplications defined above coincide and are abelian.

Proof. Item 1 is obvious and Item 2 is easily checked by verifying that the concatenation of a map with its time-reversed copy ($f(1-t, x)$) is homotopic to the constant map.

In item 3, such maps can be denoted $f_{t,s}(x)$, where t is the parameter in ΣX and s is the one in ΩY . The two multiplications differ only in which of these two parameters maps get concatenated along. The idea is then to homotopically deform one into the other. This can be done by using a deformation retract of the square $[0, 1]^2$ into a smaller square in the center. Since we're working with pointed spaces, the whole boundary of the square gets mapped to the basepoint y_0 . Hence we can replace f and g with their retractions to the smaller square, whereas outside of that square the value is just y_0 . Now the two different concatenations are clearly homotopic to each other since we can smoothly “slide” the domains of f and g around each other, and then undo the original deformation retract. This proves both the equality of the two products and their commutativity. See figure.



■

Lemma 2.5.15. For any pair of pointed spaces X, Y , there is a natural isomorphism of groups $[\Sigma X, Y] \approx [X, \Omega Y]$.

Proof. The isomorphism is induced by the correspondence that identifies $f(t, x)$ with $\gamma(x)(t) = f(t, x)$. One can easily check that this provides a one-to-one correspondence between pointed maps and pointed homotopy on both sides. Furthermore, it is almost obvious that the multiplication on both sides works the same way. ■

Definition 2.5.16 (Zeroth degree homotopy set). Denote by $\pi_0(X)$ the set of path-connected components of X . It can be identified with the set of homotopy equivalence classes of maps from the one-point space into X . Note that this set carries no natural structure of a group.

Remark 2.5.17. One of the most useful consequences of the above discussion is that in sufficiently nice topological categories, such as that of topological manifolds, homotopies are just paths in the space of continuous maps. Therefore homotopy classes of maps $[X; Y]$ are nothing but the path-connected components in the space Y^X , that is, $\pi_0(Y^X)$. For compact X Lemma 2.5.15 can be shown indirectly using the exponential law, $Y^{\Sigma X} \cong (\Omega Y)^X$. As we've just seen, the multiplications on these spaces correspond, thus we have group homomorphisms $[\Sigma X; Y] = \pi_0(Y^{\Sigma X}) \cong \pi_0((\Omega Y)^X) = [X; \Omega Y]$.

Proposition 2.5.18.

1. π_1 is a covariant functor $\pi_1 : \text{Top}_\bullet \rightarrow \text{Gr}$.
2. For a path-connected space, fundamental groups with different base points are isomorphic (this is why we often abuse the notation by writing $\pi_1(X)$).
3. For any number of path-connected spaces, $\pi_n(\prod_\alpha X_\alpha) \cong \prod_\alpha \pi_n(X_\alpha)$.

Proof. Exercise. ■

Lemma 2.5.19. If a space X is the union of a collection of path-connected open sets U_α each containing the basepoint x_0 and if each intersection $U_{\alpha\beta}$ is path-connected, then every loop in X at x_0 is homotopic to a product of loops each of which is contained in a single U_α .

Proof. The key is in the compactness of the interval $[0, 1]$, which guarantees that the loop can be broken up into a finite number segments each of whose closure is contained inside only one of the U_α 's. Call these segments $f_i, i = 1, \dots, n$. Since x_0 belongs to each U_α , we can then find a path g_i from x_0 to the starting point of f_i . We can always choose g_0 and g_{n+1} to be trivial. The loops we seek are then $g_i \cdot f_i \cdot g_{i+1}^{-1}$. Their product is homotopic to the original loop because all g_i 's cancel out. ■

Corollary 2.5.20. $\pi_1(\mathbb{S}^n) = 0$ for $n \geq 2$.

Proof. \mathbb{S}^n is the union of the open neighborhoods of two hemispheres, U_+ and U_- . Each of these is homeomorphic to \mathbb{R}^n and $U_+ \cap U_- \cong \mathbb{S}^{n-1} \times \mathbb{R}$. If $n \geq 2$, then $U_+ \cap U_-$ is path-connected. By the above lemma, every loop in \mathbb{S}^n based at a point $x_0 \in U_+ \cap U_-$ is a

product of two loops within each hemisphere. However both of these loops are trivial since $\pi_1(U_\pm) = 0$. ■

Corollary 2.5.21. \mathbb{R}^2 is not homeomorphic to \mathbb{R}^n for $n \neq 2$.

Proof. Assume there is a homeomorphism $f : \mathbb{R}^2 \rightarrow \mathbb{R}^n$. The case $n = 1$ is easy because $\mathbb{R} \setminus \{f(0)\}$ is not connected whereas the supposedly homeomorphic $\mathbb{R}^2 \setminus \{0\}$ is. When $n > 2$, the complement $\mathbb{R}^n \setminus \{f(x)\}$ is homeomorphic to $\mathbb{S}^{n-1} \times \mathbb{R}$, whose fundamental group is $\pi_1(\mathbb{S}^{n-1}) = 0$ by Proposition 2.5.18. However, for $n > 2$ the fundamental group of the complement $\mathbb{R}^2 \setminus \{x\}$ is $\pi_1(\mathbb{S}^1) = \mathbb{Z}$. This leads to a contradiction. ■

We are now prepared to prove one of the most important theorems about the fundamental group. First let us recall the categorical notion of pushouts (see Definition 1.6.4). If a path-connected space is represented as a union of two open sets, $X = U \cup V$, then the corresponding set inclusions induce homomorphisms of fundamental groups:

$$\begin{array}{ccc} X & \xleftarrow{l} & V \\ \uparrow k \quad \curvearrowright & & \uparrow j \\ U & \xleftarrow{i} & U \cap V \end{array} \quad \begin{array}{ccc} \pi_1(X) & \xleftarrow{l_*} & \pi_1(V) \\ \uparrow k_* \quad \curvearrowright & & \uparrow j_* \\ \pi_1(U) & \xleftarrow{i_*} & \pi_1(U \cap V) \end{array} \quad (2.5.3)$$

X itself is in fact the pushout, $X = U \sqcup_{U \cap V} V$. *The Seifert-van Kampen theorem will show that the functor π_1 preserves pushouts*, i.e. the fundamental group of the union is isomorphic to the amalgamated free product, $\pi_1(X) \cong \pi_1(U) *_{\pi_1(U \cap V)} \pi_1(V)$. By the universal property of the free product, there is a unique homomorphism Φ in the diagram below that makes the external triangles commute.

$$\begin{array}{ccc} \pi_1(U) * \pi_1(V) & \xleftarrow{\exists! \Phi} & \pi_1(X) \xleftarrow{l_*} \pi_1(V) \\ \nearrow & & \uparrow k_* \\ & & \pi_1(U) \xleftarrow{i_*} \pi_1(U \cap V) \end{array} \quad (2.5.4)$$

Note, however, that the “background square” involving $\pi_1(U \cap V), \pi_1(U), \pi_1(V)$, and $\pi_1(U) * \pi_1(V)$ does not commute, so the universal property of the pushout will not apply

in the other direction to give a homomorphism $\pi_1(X) \rightarrow \pi_1(U) * \pi_1(V)$. Instead, we will compute $\pi_1(X)$ by finding the kernel of Φ and factoring it out.

Theorem 2.5.22 (Siefert-van Kampen). *If X is the union of path-connected open sets U_α each containing the basepoint $x_0 \in X$ and if each intersection $U_{\alpha\beta}$ is path-connected, then the homomorphism $\Phi : *_\alpha \pi_1(U_\alpha) \rightarrow \pi_1(X)$ induced by the inclusions $U_\alpha \hookrightarrow X$ is surjective.*

Denote the inclusion-induced homomorphisms $i_\alpha : \pi_1(U_\alpha) \rightarrow \pi_1(X)$ and $i_{\alpha\beta} : \pi_1(U_{\alpha\beta}) \rightarrow \pi_1(U_\alpha)$. We have $i_\alpha \circ i_{\alpha\beta} = i_\beta \circ i_{\beta\alpha}$ since both sides are induced by the inclusion $U_{\alpha\beta} \hookrightarrow X$. Now, if in addition each intersection $U_{\alpha\beta\gamma}$ is path-connected, then the kernel of Φ is the normal subgroup N generated by all elements of the form $i_{\alpha\beta}(g)i_{\beta\alpha}(g)^{-1}$ for $g \in \pi_1(U_{\alpha\beta})$, and hence Φ induces a homomorphism $\pi_1(X) \cong *_\alpha \pi_1(U_\alpha)/N$.

In particular, in the case of two sets, $\pi_1(U \cup V) \cong \pi_1(U) *_{\pi_1(U \cap V)} \pi_1(V)$.

Proof. The surjectivity of Φ was already shown in Lemma 2.5.19.

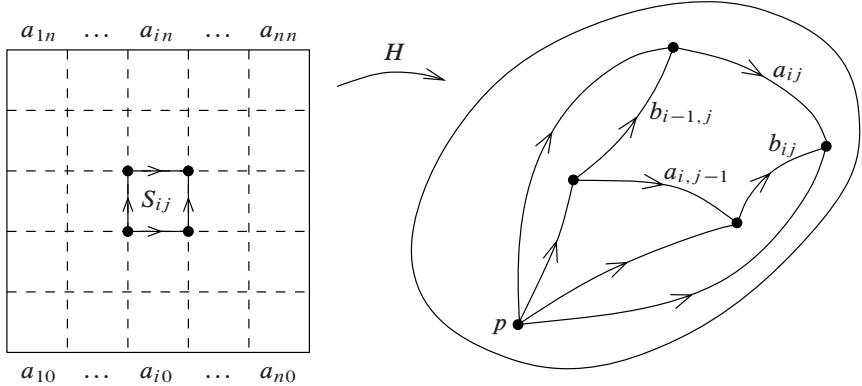
Now we need to show that the kernel of Φ is exactly N . We already know that $N \subset \ker \Phi$ since $i_\alpha \circ i_{\alpha\beta} = i_\beta \circ i_{\beta\alpha}$. For the reverse inclusion, the idea is to study *factorizations* of loops into “smaller” loops $[\gamma_i]$ that are completely contained in one of the U_α ’s. There are two “moves” that, as we would like to show, transform differing factorizations of the same loop into each other. Namely, a substring of the form $[\gamma_i][\gamma_{i+1}]$ can be replaced with $[\gamma_i \cdot \gamma_{i+1}]$, and an element $[\gamma_i]$ viewed as an element of $\pi_1(U_\alpha)$ can be replaced with its analog in $\pi_1(U_\beta)$ if γ_i is a loop contained in $U_{\alpha\beta}$. If we can show that all factorizations of a loop can be transformed into each other via a sequence of moves of this kind (or their inverses), that will mean that Φ induces an *injective* map from the quotient group into $\pi_1(X)$, and thus the kernel of Φ is exactly N .

The goal is therefore to decompose any homotopy of loops into a sequence of moves like this. A homotopy is a continuous map defined on the unit square, $H : [0, 1] \times [0, 1] \rightarrow X$, whose value on the left and right edges of the square is x_0 . Now we need to subdivide this square into smaller squares as in the Figure, so that each square is mapped entirely into only one of the open sets U_α , and so that the resulting subdivisions of the top and bottom edges of the square are refinements of the subdivisions implied in the two factorizations at hand (that is, each interval of the new subdivision is entirely contained in an interval from the old). Lastly, we can “perturb” the squares in the central rows so that each point is contained in at most three squares (this is why triple intersections were required to be path-connected).

9	10	11	12
5	6	7	8
1	2	3	4

9	10	11	12
5	6	7	8
1	2	3	4

Now we proceed to define a sequence of paths along the edges of the subdivision from the left to the right side of the square that progressively “capture” more and more of the squares. For each segment we need to specify which of the open sets in X its image is supposed to belong to, which is exactly the second move (in other words, $i_{\alpha\beta*}([\gamma_i])$ and $i_{\beta\alpha*}([\gamma_i])$ belong to the same coset in the quotient group). Further, each successive path is homotopic to the last via a deformation of a loop inside one U_α . All together, this proves that any two factorizations of the same loop are equivalent in the sense that they differ only by relations that define the group N . The illustration below is from [8, Chapter 10].



See [22, Theorem 1.20] for more details of the proof. ■

The conditions on path-connectedness of intersections and on the basepoint seem artificial, and indeed the are. Any open covering of X can be turned into a cover of this kind by taking unions of sets lying along paths connecting them to the original basepoint. The full version of the Seifert-van Kampen theorem can be stated purely in categorical terms with no extra conditions, however first we will need the notion of the *fundamental groupoid*.

Definition 2.5.23 (Groupoid). A groupoid is a set with a unary “inversion” map $x \mapsto x^{-1}$ and a partial (i.e. not necessarily defined for any pair of elements) associative binary “multiplication” operation $(x, y) \mapsto x \bullet y$. It is associative in the sense that if both $x \bullet y$ and $y \bullet z$ exist, then both $(x \bullet y) \bullet z$ and $x \bullet (y \bullet z)$ exist and coincide. It is also required that $x \bullet x^{-1}$ and $x^{-1} \bullet x$ exist, and that, if $x \bullet y$ exists, then $x \bullet y \bullet y^{-1} = x$ and $x^{-1} \bullet x \bullet y = y$.

As a consequence of these axioms, $(x^{-1})^{-1} = x$ and $(x \bullet y)^{-1} = y^{-1} \bullet x^{-1}$. Intuitively, groupoids can be thought of as groups with many identities, where only elements belonging to the same identity (which is the case if $x \bullet x^{-1} = x \bullet x^{-1}$) can be multiplied.

An equivalent definition is that a groupoid is a category where all morphisms are invertible (i.e. isomorphisms). The objects of this category are the “identities” of the groupoid. A group is then simply a groupoid with a single object. The category of groupoids is denoted Grpd .

Free products of groupoids can be constructed in the same way as for groups – as strings of “words” $g_1^{\epsilon_1} \cdots g_n^{\epsilon_n}$ where each product $g_i^{\epsilon_i} \bullet g_{i+1}^{\epsilon_{i+1}}$ must exist and insertions of substrings of the form gg^{-1} produce an equivalent string. These free products have the same universal property and are therefore the pushouts in Grpd .

Definition 2.5.24 (Fundamental groupoid). Given a topological space X , let $\Pi(X)$ (also denoted $\Pi_1(X)$) be the category whose objects are points of X and the morphisms $\text{Mor}(x, y)$ are the homotopy equivalence classes of paths $\gamma : [0, 1] \rightarrow X$ such that $\gamma(0) = x, \gamma(1) = y$. Since continuous images of homotopic paths are homotopic, we have a functor $\Pi : \text{Top} \rightarrow \text{Grpd}$.

Notice that the set of endomorphisms of any $x \in X$ viewed as an object in $\Pi(X)$ is exactly the group $\pi_1(X, x)$. Moreover, it’s not difficult to check that homotopies of continuous maps in Top induce natural transformations of the corresponding Π -functors.

Proposition 2.5.25. *For a path-connected space X , the inclusion $\pi_1(X, x) \hookrightarrow \Pi(X)$ is an equivalence of categories.*

Proof. Indeed, viewed as a category with a single object, $\pi_1(X, x)$ forms a skeleton of $\Pi(X)$, since all objects of $\Pi(X)$ are isomorphic due to path-connectedness. ■

Theorem 2.5.26 (Seifert-van Kampen). *Let $\mathcal{U} = \{U_\alpha\}_\alpha$ be an open covering of a space X that is closed under finite intersections. Regard \mathcal{U} as an index category whose morphisms are the inclusions of sets. The functor Π , restricted to sets and maps in \mathcal{U} , produces a diagram of groupoids $\Pi(\mathcal{U})$, i.e.*

$$\Pi|_{\mathcal{U}} : \mathcal{U} \rightarrow \text{Grpd}. \quad (2.5.5)$$

Then the groupoid $\Pi(X)$ is the colimit (or pushout) of this diagram (that is, the initial object in the category of cones over $\Pi|_{\mathcal{U}}$, see Section 21.1).

Proof. We can verify the universal property. For a groupoid G and a set of groupoid homomorphisms $\eta_\alpha : \Pi(U_\alpha) \rightarrow G$ (forming a cone over $\Pi|_{\mathcal{U}}$), we must construct a groupoid homomorphism $\tilde{\eta} : \Pi(X) \rightarrow G$ (which is nothing but a functor between categories) that restricts to η_α on $\Pi(U_\alpha)$ for all α . On objects, that is on points of X , we have no choice but to define $\tilde{\eta}(x) = \eta_\alpha(x)$ for $x \in U_\alpha$. This is independent of the choice of α since the cover is closed under finite intersections.

Now we must define $\tilde{\eta}$ on morphisms in $\Pi(X)$, i.e. paths. If a path γ connecting x to y lies entirely in a particular U_α , we again have only one choice: $\tilde{\eta}([\gamma]) = \eta_\alpha([\gamma])$. This is consistent (independent of α) since \mathcal{U} is closed under finite intersections.

Now, any path γ is the composite of finitely many paths γ_i , each of which lies in a single U_{α_i} . Therefore we must define $\tilde{\eta}([\gamma]) = \tilde{\eta}([\gamma_1]) \bullet \dots \bullet \tilde{\eta}([\gamma_n])$.

Clearly this definition will give the required unique morphism $\tilde{\eta}$, provided that it is indeed well defined, i.e. respects the equivalence relation between paths. Thus suppose we have two equivalent paths $\gamma_0 \sim \gamma_1$. The equivalence is given by a homotopy γ_t through paths with the same fixed endpoints. This homotopy is a map defined on the square $[0, 1] \times [0, 1]$. We may subdivide the square into subsquares, each of which is mapped entirely into one of the U_α 's. We may further choose this subdivision so that the resulting subdivision of $[0, 1] \times \{0\}$ refines the subdivision used to decompose γ_0 into segments as above, and similarly for γ_1 and the resulting decomposition of $[0, 1] \times \{1\}$. We see that the relation $[\gamma_0] = [\gamma_1]$ in $\Pi(X)$ is a consequence of a finite number of relations, each of which holds in one of the $\Pi(U_\alpha)$'s. Therefore $\tilde{\eta}([\gamma_0]) = \tilde{\eta}([\gamma_1])$. This confirms that $\tilde{\eta}$ is well defined and satisfies the universal property for pushouts. ■

From this, one can use an “abstract nonsense” argument to re-derive the original version of the theorem for the functor $\pi_1(-, x_0)$, see [31, §2.7].

Corollary 2.5.27 (Fundamental group of a wedge sum). *Suppose x_α is a nondegenerate basepoint for X_α (which means that x_α has a neighborhood which admits a strong deformation retraction into x_α). Then the basepoint of $\vee_\alpha X_\alpha$ is nondegenerate and the fundamental group of this wedge sum is isomorphic to the free product $*_\alpha \pi_\alpha(X_\alpha)$.*

Proof. See [8, Theorem 10.7]. ■

Example 2.5.28. 1. The shape Θ is a deformation retract of the figure 8, which itself is just the wedge sum $\mathbb{S}^1 \vee \mathbb{S}^1$, so their fundamental groups are all $\mathbb{Z} * \mathbb{Z}$.
 2. The last example can be generalized to compute π_1 of an arbitrary connected graph. The

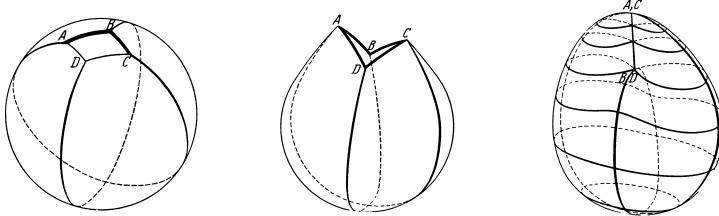


Fig. 2.2: Construction of the projective plane out of a disk. Adding a hole to this surface creates a crosscap, which is homeomorphic to a Möbius band, see Exercises 2.20 and 2.21.

result is just the free group generated by the set of “elementary” cycles in the graph. More precisely, the number of generators equals the number of all edges of the graph minus the number of edges in a maximal spanning tree of the graph (i.e. a maximal subgraph without cycles).

Exercise 2.18. [8, Exercise 10-1] Use the Seifert-van Kampen theorem to give another proof that \mathbb{S}^n is simply connected when $n \geq 2$.

Exercise 2.19. [8, Exercise 10-5] Compute the fundamental group of \mathbb{R}^3 with the three coordinate axes removed. *Hint:* this space is homotopy equivalent to the 2-sphere with six points removed.

Exercise 2.20 (Projective plane). Show that the projective plane \mathbb{RP}^2 , constructed out of a disk by identifying the antipodal points on the bounding circle (see Figure 2.2), has the fundamental group \mathbb{Z}_2 .

Exercise 2.21 (Crosscap). Show that the projective plane with a hole (also known as a *crosscap*), constructed out of an annulus by identifying the antipodal points of the outer boundary, is homeomorphic to the Möbius band, and therefore has the fundamental group \mathbb{Z} .

Example 2.5.29. [19, Exercise III.3] Consider an annulus. Identify antipodal points on the outer circle. Also identify antipodal points on the inner circle. Calculate the fundamental group of this surface. Is it homeomorphic to any of the orientable compact surfaces? *Hint:* taking U and V to be overlapping annular neighborhoods of the two boundaries of the annulus, the generator of $\pi_1(U \cap V)$ is mapped to the square of the generator in $\pi_1(U)$ and similarly in $\pi_1(V)$; therefore $\pi_1(U \cup V) = \langle a, b \mid a^2 = b^2 \rangle$.

Exercise 2.22 ([22, Exercise 1.20]). Let X be the subspace of \mathbb{R}^2 that is the union of the circles C_n of radius n and center $(n, 0)$ for $n = 1, 2, \dots$. Show that $\pi_1(X)$ is the free group $*_n \pi_1(C_n)$, the same as for the infinite wedge sum $\bigvee_{i=1}^{\infty} \mathbb{S}^1$. Show that X and $\bigvee_{i=1}^{\infty} \mathbb{S}^1$ are in

fact homotopy equivalent, but not homeomorphic.

Example 2.5.30 (Hawaiian earring [22, Example 1.25]). Let X be the subspace of \mathbb{R}^2 that is the union of the circles C_n of radius $1/n$ and center $(1/n, 0)$ for $n = 1, 2, \dots$. Initially this space can be mistaken for the infinite wedge sum $\bigvee_{i=1}^{\infty} \mathbb{S}^1$, but in fact it has a much larger fundamental group. This is because in the topology induced from \mathbb{R}^2 it is possible for a continuous loop $\gamma[0, 1] \rightarrow X$ to cover an infinite number of circles C_n at once, unlike in the case of the wedge sum. As a result, $\pi_1(X)$ can be mapped surjectively onto the infinite direct product group $\prod_{n=1}^{\infty} \mathbb{Z}$, which is uncountably large.

Example 2.5.31 (Torus knots [22, Example 1.24]). For relatively prime integers m, n , the *torus knot* $K = K_{m,n} \subset \mathbb{R}^3$ is the image of the embedding $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1 \times \mathbb{S}^1 \cong \mathbb{T}^2 \subset \mathbb{R}^3$ given by $f(z) = (z^m, z^n)$ ($z = e^{2i\pi t}$), where the torus T is embedded in space in the standard way. Let us compute $\pi_1(\mathbb{R}^3 \setminus K)$. First, we can safely replace \mathbb{R}^3 with its compactification \mathbb{S}^3 by adding “a point at infinity”. This isomorphism follows from the Seifert-van Kampen theorem. Now we may decompose $\mathbb{S}^3 \setminus \mathbb{T}^2$ into two open solid tori U, V . Let C be an open tubular neighborhood of $T \setminus K$. Then $U \cup C$ and $V \cup C$ form an open covering of $\mathbb{S}^3 \setminus K$ with a path connected intersection. The fundamental groups of $U \cup C, V \cup C$, and the intersection C are all isomorphic to \mathbb{Z} and generated by loops x, y, z , respectively. The inclusions of C into $V \cup C$ or $U \cup C$ map z to x^m or y^n . Therefore $\pi_1(\mathbb{S}^3 \setminus K) = \langle x, y \mid x^m = y^n \rangle$. The center of this group is the subgroup generated by the element $x^m = y^n$, and the factor by the center is exactly $\mathbb{Z}_m * \mathbb{Z}_n$. By commuting all powers of x^m through to the left, the elements of this group can be uniquely represented by strings of the form $x^{mk} y^{l_0} x^{k_1} y^{l_1} x^{k_2} y^{l_2} \dots$ where $k \in \mathbb{Z}$, $k_i \in 0, \dots, m$ and $l_i \in \{0, \dots, n\}$.

Example 2.5.32 (Lens spaces [19, Exercise III.10]). Let $X = (\mathbb{D}^2 \times \mathbb{S}^1) \sqcup_f (\mathbb{D}^2 \times \mathbb{S}^1)$ be the gluing of two solid tori along the map $f : \mathbb{S}^1 \times \mathbb{S}^1 \rightarrow \mathbb{S}^1 \times \mathbb{S}^1$ induced from the linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by an integer matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The fundamental group of a solid torus are generated by one generator, call it γ_1 and γ_2 for the two solid tori correspondingly; in addition, the two solid tori overlap by a space that deformation retracts onto a torus \mathbb{T}^2 whose fundamental group is generated by two simple generators α and β ; assuming $a \neq 0$, by Seifert-van Kampen, from one inclusion we have the mappings $\alpha \mapsto \gamma_1$ and $\beta \mapsto e$, whereas from the other we get $\alpha \mapsto \gamma_2^d$ and $\beta \mapsto \gamma_2^c$. The resulting identification $\gamma_1 \sim \gamma_2^d$ simply allows us to drop γ_1 from the list of generators. Thus $\pi_1(X)$ is the cyclic group $\langle \gamma_2 \mid \gamma_2^c \rangle = \mathbb{Z}_c$. Note that the case $a = 0$ is special – in that case $\pi_1(X) = \mathbb{Z}$. If the matrix of f is invertible (which is true iff $\det f = \pm 1$) and is parametrized as $\begin{pmatrix} q & (qq' - 1)/p \\ p & q' \end{pmatrix}$, the resulting space is the so called *lens space* $L(p; q)$ with fundamental group \mathbb{Z}_p . We have $L(1; 0) = \mathbb{S}^3$ ($f = \text{id}$) and $L(2; 1) = \mathbb{RP}^3$ ($f = -\text{id}$).

3 Manifolds

3.1 Topological manifolds

Definition 3.1.1 (Locally Euclidean Space). A topological space X is locally n -dimensional Euclidean if every $x \in X$ has an open neighborhood homeomorphic to the open n -dimensional ball, $U_x \cong B^n$.

Definition 3.1.2 (Second-countability). A topological space is called second countable if it has a countable basis of topology.

Definition 3.1.3 (Paracompactness).

1. A collection of subsets $A_\alpha, \alpha \in I$, of X is called locally finite if every $x \in X$ has an open neighborhood U_x such that $U_x \cap A_\alpha \neq \emptyset$ only for a finite number of α 's. If all A_α 's are open, this is the same as x being an element of only finitely many of them.
2. Given an open covering $\{U_\alpha\}$ (i.e. a collection of open sets such that $\bigcup_\alpha U_\alpha = X$), a refinement of it is an open covering $\{V_\beta\}$ such that $\forall \beta \exists \alpha : V_\beta \subset U_\alpha$.
3. X is called paracompact if every open covering of X admits an open, locally finite refinement.

Definition 3.1.4 (Topological manifold). A topological manifold is a second countable, Hausdorff, locally Euclidean space.

The following proposition is the reason why second countability is often substituted with paracompactness in the definition of manifolds.

Proposition 3.1.5. *All topological manifolds are paracompact (and locally compact).*

Proof. See [7, Thm. 1.15] ■

Definition 3.1.6 (Charts). Let X be an n -dimensional topological manifold. A *chart* on it is a pair (U, φ) , where $U \subset X$ is an open set and $\varphi : U \rightarrow \mathbb{R}^n$ is a continuous map (called the chart map) that is also a homeomorphism onto its image $\hat{U} = \varphi(U)$, i.e. $U \cong \varphi(U) = \hat{U}$. The chart is called a *coordinate ball* if \hat{U} is a ball in \mathbb{R}^n . The values of φ are called *coordinates* of the corresponding points on X .

Remark 3.1.7. Since X is locally Euclidean, every point x automatically has a chart (U_x, φ) around it. In particular, every manifold has an open covering by charts $(U_\alpha, \varphi_\alpha)$.

Theorem 3.1.8. *All topological manifolds have the following properties.*

1. *There is a countable basis consisting of coordinate balls (it can also be made locally finite and every ball can be made precompact)*
2. *Local compactness.*

3. Local path-connectedness (i.e. every point has a path-connected neighborhood).
4. At most countably many connected components, each open and a topological manifold of its own.
5. The fundamental group $\pi_1(X)$ is at most countable.

Proof. 1,2) Pick an open covering by charts $\{U_\alpha\}$. Each U_α can be covered by countably many coordinate balls V_{α_i} because \bar{U}_α can be covered by countably many balls in \mathbb{R}^n . By second countability, there is a countable subcovering and basis $\{V_\beta\}$. Fixing a chart U and covering it by coordinate balls V_α , we have that \bar{V}_β is compact in U (because it's homeo to a closed ball), therefore closed in X by Hausdorffness. Therefore it's also compact in X , i.e. V_β 's are precompact. The covering can also be made locally finite: let $\{K_i\}_{i=1}^\infty$ be an exhaustion of M by compact sets (i.e. $\bigcup_i K_i = M$; we will prove the existence in Prop. 3.3.6). After picking out a finite subcovering of every compact “stripe” $K_{i+1} \setminus \text{int } K_i$ by charts that are small enough to fit within the wider open “stripe” $\text{int } K_{i+2} \setminus K_{i-1}$ (possible because all charts still form a basis), we end up with a locally finite covering (since every point of the manifold is contained only in 3 of the open stripes).

3) Obvious because balls are.

4) 2nd countability implies countable components. The rest is obvious.

5) See [7, Thm 1.16]. ■

Lemma 3.1.9. Locally path-connected and connected \Leftrightarrow path-connected.

Proof. The reverse direction is obvious. In the forward direction, introduce the set $C = \{q : q \text{ is connected with } p\}$ for some fixed $p \in X$. We will prove that C is clopen and therefore equal to X .

C is open because X is locally path-connected. C is closed because for any $\underline{z} \in \bar{C}$ there is a path-connected neighborhood U_z , thus $U_z \cap C \neq \emptyset$, and $\underline{z} \in C$. Therefore $\bar{C} = C$. But X is connected, and C is clearly not empty, therefore $C = X$, and lastly C is path-connected by definition. ■

Corollary 3.1.10. For manifolds, connectedness is equivalent to path-connectedness. In particular, connected components of manifolds are path-connected.

Example 3.1.11. 1. Any open set in \mathbb{R}^n is a manifold.

2. \mathbb{S}^n is a manifold because it can be covered by finitely many charts (U_i, φ_i) , where U_i is an open hemisphere whose “equator” lies on one of the coordinate hyperplanes, and φ_i is simply the orthogonal projection to that hyperplane.
3. An n -dimensional torus $\mathbb{T}^n = (\mathbb{S}^1)^{\times n}$ is a manifold.
4. The figure eight, 8, as a subset of \mathbb{R}^2 , is not a manifold, because its center doesn't have a

neighborhood homeomorphic to an interval.

5. Matrices $\text{Mat}(n, \mathbb{R})$ form an n^2 -dimensional manifold (isomorphic to \mathbb{R}^{n^2}). So do invertible matrices $\text{GL}(n, \mathbb{R})$ (being an open subset in \mathbb{R}^{n^2}) and all other classical matrix groups.

3.2 Smooth structures

Definition 3.2.1 (Atlas). An atlas \mathcal{A} for a topological manifold M is a collection of continuous charts $\{(U_\alpha, \varphi_\alpha)\}$, $\varphi_\alpha : U_\alpha \rightarrow \hat{U}_\alpha \subset \mathbb{R}^n$ such that $\bigcup_\alpha U_\alpha = M$.

Throughout these notes we also adopt the notation where for any atlas as above, $U_{\alpha\beta} = U_\alpha \cap U_\beta$, $U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma$, etc.

Definition 3.2.2 (Smooth atlas). An atlas is called smooth (or of class C^k or C^ω) if all of the *transition functions* $\varphi_{\beta\alpha} = \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_{\alpha\beta}) \rightarrow \varphi_\beta(U_{\alpha\beta})$ are smooth/ C^∞ (of class C^k , or real analytic, respectively) in the sense of multivariable calculus on \mathbb{R}^n .

Definition 3.2.3 (Equivalent atlases). Two atlases \mathcal{A}_1 and \mathcal{A}_2 are called equivalent, or compatible, if $\mathcal{A}_1 \cup \mathcal{A}_2$ is still an atlas of the same smoothness class.

Definition 3.2.4 (Differentiable structure). A smooth (C^k , analytic) structure on a topological manifold M is an equivalence class $[\mathcal{A}]$ of smooth (respectively, C^k , C^ω) atlases on it.

Definition 3.2.5 (Smooth manifold). A smooth manifold is a pair $(M, [\mathcal{A}])$ consisting of a topological manifold and a smooth structure.

Remark 3.2.6. Recall that the second countability of topological manifolds implies their paracompactness (but not vice versa). In the case of differentiable manifolds, these properties turn out to be equivalent, and also equivalent to metrizability, or to the existence of smooth partitions of unity (see next section). This is why the definitions sometimes differ.

Example 3.2.7. 1. Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be $\psi(x) = x^3$. Then $\{(\mathbb{R}, \psi)\}$ is an atlas that defines a smooth structure on \mathbb{R} . It is not smoothly compatible with the standard smooth structure given by the chart map $\varphi(x) = x$, because the transition function $\varphi \circ \psi^{-1}(x) = x^{1/3}$ is not smooth. Therefore \mathbb{R} has many (in fact, infinitely many) different smooth structures.
2. Any real or complex vector space is automatically a smooth manifold.
3. Spaces of matrices, classical groups, spheres... are all smooth manifolds with their standard structures.
4. Not every topological manifold admits a smooth structure, although examples are pretty difficult to construct. On the other hand, any C^1 manifold admits a compatible (in the C^1 sense) smooth structure, which is unique up to diffeomorphism (see below). This is why even mathematicians are mostly happy working with just either topological

or immediately smooth/analytic manifolds. Interestingly, all topological manifolds in dimensions 1,2, and 3 admit unique (up to diffeo) smooth structures.

5. All Lie groups (defined as smooth manifolds with a group structure in which multiplication and inversion are smooth maps) in fact support a unique (up to diffeo) C^ω structure. This is an immediate consequence of the convergence of the Baker-Campbell-Hausdorff series.

From now on we work only with smooth manifolds, and use the word “manifold” to mean “smooth manifold”.

Definition 3.2.8 (Smooth maps). Let M and N be two manifolds with atlases $\{(U_\alpha, \varphi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$, respectively. A map $f : M \rightarrow N$ is called smooth if all of its *local representatives* $f_{\beta\alpha} = \psi_\alpha \circ f \circ \varphi_\beta^{-1} : \hat{U}_\beta \rightarrow \hat{U}_\alpha$ are smooth. The set of all smooth maps between two manifolds is denoted by $C^\infty(M, N)$. When $N = \mathbb{R}$, we write $C^\infty(M) = C^\infty(M, \mathbb{R})$ for the space of “smooth functions”. The category Man^∞ is the category of smooth manifolds with smooth maps for morphisms.

Exercise 3.1. Check that all chart maps in any atlas of a smooth manifold M are smooth as maps from M to the standard \mathbb{R}^n .

Remark 3.2.9. The above exercise illustrates the general alternative approach to manifolds: instead of defining a smooth structure in terms of charts, we could describe it by specifying *what smooth functions should be* by describing the subsets of the sets of local continuous functions $C(U_\alpha)$ that we want to be smooth (more precisely, one has to pick a subsheaf of the sheaf of continuous functions that is locally isomorphic to the sheaf of smooth functions on \mathbb{R}^n).

Definition 3.2.10 (Manifolds with boundaries). Let \mathbb{R}_+^n be the closed upper half-space in \mathbb{R}^n (i.e. $x^n \geq 0$). An n -dimensional manifold with boundaries is defined just like a smooth manifold, only now some of the chart maps can map to \mathbb{R}_+^n . The smoothness of boundary transitions is characterized as follows: any function $\mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$ is called smooth if it can be extended as a smooth function into an open neighborhood of \mathbb{R}_+^n . The set of points of M that map to the boundary of $\mathbb{R}_+^n \subset \mathbb{R}^n$, i.e. the hyperplane $x^n = 0$, is called the boundary ∂M . We also denote the interior by $\mathring{M} = M \setminus \partial M$.

Example 3.2.11. $\sqrt{y} : \mathbb{R}_+ \rightarrow \mathbb{R}$ is not smooth on \mathbb{R}_+ with the standard smooth structure. However, it can be taken to be the chart map defining a new smooth structure on \mathbb{R}_+ , and in that new structure it'll become smooth, because its local representative will be the identity map.

Remark 3.2.12. Similarly, one can define *manifolds with corners* using chart maps that can take values in quadrants, octants, and so on, of \mathbb{R}^n .

Definition 3.2.13 (Diffeomorphisms). A smooth map $f : M \rightarrow N$ between manifolds is called a diffeomorphism if it has a smooth inverse. They are the isomorphisms in Man^∞ .

Proposition 3.2.14. 1. Diffeos are homeo and open.

2. If M and N have boundaries and $f : M \rightarrow N$ is a diffeo, then $f(\partial M) = \partial N$ and $f|_{\overset{\circ}{M}} : \overset{\circ}{M} \rightarrow \overset{\circ}{N}$ is a diffeo.

Example 3.2.15. 1. Different smooth structures can be diffeomorphic! Take \mathbb{R} with the chart $\psi(x) = x^3$ that we considered earlier. It is diffeomorphic to the standard smooth structure given by the chart $\varphi(x) = x$, and the diffeomorphism is given by ψ itself! Indeed, its local representative in the two charts is $\varphi \circ \psi \circ \psi^{-1}(x) = x$. In this manner one can show that there is only one smooth structure on \mathbb{R} up to diffeomorphism.

2. Historically the first manifolds to be found to support more than one non-diffeomorphic smooth structure were spheres, starting with \mathbb{S}^7 which has 28 diffeomorphism classes of smooth structures. These smooth structures are now known as Milnor's *exotic spheres*.
3. \mathbb{R}^n has only one smooth structure up to isomorphism for all $n \neq 4$. For $n = 4$, it admits *uncountably many* non-diffeomorphic structures. This makes the study of four-dimensional topological manifolds especially sophisticated, and there are still many very fundamental unsolved problems in four-dimensional topology [12].
4. In every dimension $n \geq 4$ there exist *compact* topological manifolds that don't admit any compatible smooth structure.

Exercise 3.2. Show that \mathbb{S}^1 supports only one smooth structure up to diffeomorphism.

3.3 Partitions of unity

First we note that on \mathbb{R}^n there exist bump functions, i.e. for any point x_0 and any two radii $0 < r_1 < r_2$ there is a function $\eta \in C^\infty(\mathbb{R}^n)$ such that $\eta|_{B_{r_1}(x_0)} = 1$ and $\eta|_{\mathbb{R}^n \setminus \overline{B_{r_2}}(x_0)} = 0$.

Exercise 3.3. Using the function $\int_a^x e^{\frac{1}{(x'-a)(b-x')}} dx'$ defined on (a, b) , construct a general bump function for \mathbb{R}^n .

Now, any manifold M can be covered by a locally finite collection of coordinate balls $(U_\alpha, \varphi_\alpha)$. Let η_α be a bump function that equals 0 outside of the ball $\hat{U}_\alpha \subset \mathbb{R}^n$. Then define the smooth functions

$$f_\alpha = \begin{cases} \eta_\alpha \circ \varphi_\alpha, & \text{on } U_\alpha, \\ 0, & \text{on } M \setminus U_\alpha. \end{cases} \quad (3.3.1)$$

Finally, we normalize this collection of functions by defining

$$\chi_\alpha(x) = \frac{f_\alpha(x)}{\sum_\alpha f_\alpha(x)}. \quad (3.3.2)$$

This expression is well-defined because the sum in the denominator is finite for every fixed x . These functions have the following properties:

1. $\chi_\alpha \in C^\infty(M)$,
2. $0 \leq \chi_\alpha \leq 1$,
3. $\text{supp } \chi_\alpha \subset U_\alpha$,
4. $\{\chi_\alpha\}$ is locally finite, i.e. only finitely many of them are non-zero at any given point $x \in M$,
5. $\sum_\alpha \chi_\alpha(x) \equiv 1$.

Definition 3.3.1 (Partition of unity). Any collection of functions satisfying the above list of properties is called a **partition of unity (POU)** subordinate to the open covering (not necessarily by coordinate balls) $\{U_\alpha\}$ of M .

Proposition 3.3.2 ([8, Thm. 4.85]). *Any open covering of a paracompact Hausdorff space admits a subordinate POU.*

Proposition 3.3.3. *Any open covering of a smooth manifold admits a subordinate partition of unity.*

Proof. Any open covering has a locally finite refinement consisting of coordinate balls, which allows us to apply the construction above to the refinement. Notice that second countability is key to the existence of partitions of unity. ■

Theorem 3.3.4 (Extension lemma). *Let $A \subset M$ be a closed subset, and $f \in C^\infty(A, \mathbb{R}^k)$. Then for any open U such that $A \subset U$, there exists an extension $\tilde{f} \in C^\infty(M, \mathbb{R}^k)$: $\tilde{f}|_A = f$, such that $\text{supp } \tilde{f} \subset U$.*

Proof. Pick $p \in A$ and neighborhood $U_p \subset U$. Let $\tilde{f}_p : U_p \rightarrow \mathbb{R}^k$ be a smooth extension of f from $U_p \cap A$. It is guaranteed to exist for sufficiently small U_p by definition of smoothness on closed subsets of \mathbb{R}^n and smoothness of chart maps.

Then $\{U_p\}_{p \in A} \cup \{M \setminus A\}$ is an open covering of M . Let $\{\chi_p\} \cup \{\chi_0\}$ be a **POU** subordinate to this covering (and $\chi_p = 0$ inevitably for almost all p so that the resulting collection is locally finite). Therefore $\chi_p \cdot \tilde{f}_p : U_p \rightarrow \mathbb{R}^k$ are smooth and vanishing at the boundary of U_p , which allows us to extend them by zeros to smooth functions on all M . Finally, $\tilde{f} = \sum_{p \in A} \chi_p \tilde{f}_p$ is the extension we seek with support inside U . ■

Definition 3.3.5 (Exhaustion function). A function $f \in C^\infty(M)$ is called an exhaustion function for M if for every $c \in \mathbb{R}$ the pre-image $f^{-1}((-\infty, c])$ is relatively compact (i.e. has compact closure in M).

Proposition 3.3.6. *Every smooth manifold has an exhaustion function.*

Proof. Let $\{V_\alpha\}$ be a countable open covering by precompact sets and $\{\chi_i\}_{i=1}^\infty$ a subordinate POU. Define $f(x) = \sum_{n=1}^\infty n \cdot \chi_n(x)$.

Clearly $f^{-1}((-\infty, c]) \subset \bigcup_{i=1}^{[c]+1} \overline{V}_i$, which is a compact set. ■

3.4 Sard's theorem

In \mathbb{R}^n , a set A is said to have *zero measure* if for any $\epsilon > 0$ it can be covered by a collection of open balls B_i such that the total volume of these balls is less than ϵ .

Definition 3.4.1 (Measure zero). Even though there is no canonical way to measure volumes on manifolds, there is still a consistent notion of zero measure. Namely, a set $A \subset M$ has zero measure if all the sets $\varphi_\alpha(A \cap U_\alpha)$ are measure zero in \mathbb{R}^n for an atlas $\{(U_\alpha, \varphi_\alpha)\}$.

Definition 3.4.2 (Rank of a map). Let $f \in C^\infty(M^m, N^n)$ (superscripts here indicate the dimensions of the manifolds). We say that the rank of f at $p \in M$ is $\text{rank}_p f = r$ if there exist two local charts, (U, φ) around $p \in M$ and (V, ψ) around $f(p) \in N$, such that the local representative of f in these charts is $\psi \circ f \circ \varphi^{-1}(x^1, \dots, x^m) = (\underbrace{x^1, \dots, x^r}_{n-r}, 0, \dots, 0)$.

Definition 3.4.3 (Critical and regular points/values). $p \in M$ is said to be a critical point of $f \in C^\infty(M^m, N^n)$ if $\text{rank}_p f < n$, i.e. if its differential at p is not surjective. In this case, $f(p) \in N$ is called a critical value of f . Otherwise, p is a regular point, and a regular value is one all of whose pre-images are regular points.

The main property of regular values of smooth maps is that they are generic, as demonstrated by the following deep theorem. We will need this result when we talk about transversality in the context of fiber bundles.

Theorem 3.4.4 (Sard). *For any $f \in C^\infty(N, M)$ the set of critical values (also known in some contexts as the caustic) of f has measure zero in M .*

Proof. This theorem is essentially entirely local, i.e. it suffices to prove it for $N = \mathbb{R}^n$ and $M = \mathbb{R}^m$. Every manifold can be broken up into a countable number of neighborhoods, and a countable union of sets of measure zero is still measure zero.

Let $C \subset \mathbb{R}^m$ be the set of critical points of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and pick $x_0 \in C$. Then, after some rotation of the space, we can assume that its derivative maps $f'_{x_0}(\mathbb{R}^n) = \mathbb{R}^k \times \{0\} \subset \mathbb{R}^m$, where $k \leq n - 1$ is the rank of the Jacobi matrix at x_0 . If we choose a ball $B_r(x_0)$ of radius r around x_0 , then

$$f'_{x_0}(B_r(x_0)) \subset (\mathbb{R}^k \times \{0\}) \cap B_\epsilon(f(x_0)), \quad \epsilon = r \cdot \text{Lip } f, \quad (3.4.1)$$

where $\text{Lip } f$ is the Lipschitz constant of f inside $B_r(x_0)$ (equal to the supremum of all directional derivatives in unit directions). We only need the existence of such an ϵ and that it goes to zero with $r \rightarrow 0$.

Therefore $f'_{x_0}(B_r(x_0))$ can be covered by N^k cubes of diameter ϵ/N . By differentiability there exists $\delta > 0$ such that

$$|f(y) - f(x_0) - f'_{x_0}(y - x_0)| \leq \frac{1}{N} |y - x_0|, \quad \text{for all } y \in B_\delta(x_0). \quad (3.4.2)$$

So $f(y)$ is within r/N distance of the hyperplane $f(x_0) + f'_{x_0}(\mathbb{R}^n)$ (provided that $r < \delta$). Therefore the image $f(B_r(x_0))$ itself is contained in a neighborhood of $f'_{x_0}(B_r(x_0))$ of radius r/N . Thus we can construct N^k disjoint sets E_j of diameter $\sim r/N$ that cover $f(B_r(x_0))$.

We repeat the above for every point $x \in C$ and get $r_x > 0$ such that $f(B_{r_x}(x)) \subset \bigcup_{j=1}^{N^k} E_j^x$. By [Vitali's covering lemma](#), there is a countable subcovering, so $C \subset \bigcup_{i=1}^{\infty} B_{r_i}(x_i)$ such that $\{B_{r_i/5}(x_i)\}$ are disjoint and for every i there are sets E_j^i , $j = 1, \dots, N^k$ such that $\text{diam}(E_j^i) \sim r/N$ and $f(B_{r_i}(x_i)) \subset \bigcup_{j=1}^{N^k} E_j^i$. Thus $f(C) \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{N^k} E_j^i$. We can safely assume that $C \subset [-l, l]^n$ for some $l \in \mathbb{N}$ (otherwise it can be broken up into countably many pieces of this kind, which won't upset the result of the theorem). By compactness we can then assume that $r_x < 1$ for all $x \in C$. Thus

$$\begin{aligned} \text{volume}(f(C)) &\leq \sum_{i,j} (\text{diam}(E_j^i))^n \leq \sum_{i,j} \left(\frac{r}{N}\right)^n \lesssim \sum_i N^k \frac{r^n}{N^n} \lesssim \\ &\lesssim 5^n N^{k-n} \sum_i \frac{r^n}{5^n} \lesssim 2 \cdot 5^n N^{k-n} \cdot \text{volume}([-l, l]^n), \end{aligned} \quad (3.4.3)$$

since C lies inside $[-l, l]^n$ and is covered by balls of radii $r_i \leq 1$ with disjoint sub-balls, so the sum of the volumes of these sub-balls can only exceed the volume of $[-l, l]^n$ by at most a constant factor (coming from pieces that stick out of $[-l, l]^n$), say 2.

But now, $k < n$, thus in the limit $N \rightarrow \infty$ we find that the volume of $f(C)$ has to be zero. This proof is due to Behnam Esmayli, a more conventional one can be found in [7, Thm. 6.10]. ■

Remark 3.4.5. The ability to globalize the proof of this theorem depends crucially on second countability. Indeed, we could consider a non-second-countable 1-dimensional “manifold” given by the uncountable disjoint union of lines $M = \bigsqcup_{a \in \mathbb{R}} R_a = \bigsqcup_{a \in \mathbb{R}} \mathbb{R}$, and define a function $f : M \rightarrow \mathbb{R}$ by $f|_{R_a} = a$. In this case f is smooth and every value $a \in \mathbb{R}$ is a critical value, contrary to Sard’s theorem. What happened was that we found an uncountable family of local neighborhoods (R_a) , for each of which the set of critical values has measure zero (one point a). But the union of an uncountable number of sets of measure zero can have

nonzero measure. Thus second countability is key to Sard's theorem working on general manifolds.

3.5 Inverse & Implicit Mapping Theorems

In this section we just recall the statements of two basic theorems in multivariable calculus.

Theorem 3.5.1 (Inverse Mapping Theorem (InMT)). Let $U, V \subset \mathbb{R}^n$ be open and $f \in C^\infty(U, V)$. If the Jacobi matrix $\frac{\partial f^i}{\partial x^j}(a)$ is invertible for some $a \in U$, then there are neighborhoods $a \in U_a \subset U$ and $f(a) \in V_{f(a)} \subset V$ such that $f|_{U_a} : U_a \rightarrow V_{f(a)}$ is a diffeomorphism.

Meaning: any smooth map whose Jacobi matrix is non-singular at a point is locally a diffeomorphism around that point.

Theorem 3.5.2 (Implicit Mapping Theorem (ImMT)). Let $U \subset \mathbb{R}^n \times \mathbb{R}^k$ be open, $(x, y) = (x^1, \dots, x^n, y^1, \dots, y^k) \in U$. Let $\Phi \in C^\infty(U, \mathbb{R}^k)$ and $\Phi(x, y) = c$. If the matrix $\left(\frac{\partial \Phi^i}{\partial y^j}(x, y)\right)_{i,j=1}^k$ is non-singular, then there are open neighborhoods $x \in V_x \subset \mathbb{R}^n$ and $y \in W_y \subset \mathbb{R}^k$ and a smooth map $f \in C^\infty(V_x, W_y)$ such that

$$\forall (v, w) \in V_x \times W_y, \quad \Phi(v, w) = c \Leftrightarrow w = f(v). \quad (3.5.1)$$

Meaning: the equation $\Phi(x, y) = c$ can be locally solved for $y = y(x)$ if the matrix $\partial \vec{\Phi} / \partial \vec{y}$ is invertible.

3.6 Tangent spaces

In \mathbb{R}^n , vectors don't have "base points" because they can always be parallel transported to the origin. On a manifold, this is impossible, which is why there is a tangent space "attached" to every point $p \in M$.

Definition 3.6.1 (Tangent vector). Let $p \in M$. If $p \in \mathring{M}$, consider all curves $\gamma \in C^\infty((-1, 1), M)$ such that $\gamma(0) = p$. If $p \in \partial M$, consider instead all curves with domains $(-1, 0]$ or $[0, 1)$. Introduce the equivalence relation $\gamma_1 \sim \gamma_2$ iff $\frac{d}{dt}|_{t=0} (\varphi_\alpha \circ \gamma_1)(t) = \frac{d}{dt}|_{t=0} (\varphi_\alpha \circ \gamma_2)(t)$ for any chart φ_α around p . A tangent vector at p is an equivalence class $[\gamma]_p$ of curves passing through p . The set of all tangent vectors at p is denoted by $T_p M$.

1. $T_x M$ is imbued with the structure of a real vector space as follows:

- (a) $\lambda \cdot [\gamma(t)]_p := [\gamma(\lambda t)]_p$ for $\lambda \in \mathbb{R}$.
- (b) $[\gamma_1(t)]_p + [\gamma_2(t)]_p := [\varphi_\alpha^{-1}(\varphi_\alpha(\gamma_1(t)) + \varphi_\alpha(\gamma_2(t)))]_p$.

2. We will show later that $\dim T_p M = \dim M$.

Definition 3.6.2 (Tangent vector, alternative def.). $T_p M$ can be alternatively defined as the set of all *derivations* at p , i.e. linear functionals $v : C^\infty(M) \rightarrow \mathbb{R}$ such that $v[f \cdot g] = f(p)v[g] + g(p)v[f]$. This is obviously a vector space. Notably, this definition doesn't need to be modified if $p \in \partial M$.

Proposition 3.6.3. *The two definitions of tangent spaces given above are isomorphic as vector spaces.*

Proof. An explicit linear isomorphism between the two constructions is $[\gamma]_p \mapsto v[f] = \frac{d}{dt} \Big|_{t=0} f(\gamma(t))$ (this is called the *directional derivative* along the tangent vector $[\gamma]_p$).

All derivations on the algebra of smooth functions can be shown to be directional derivatives, which allows us to invert this mapping. Namely, for a chart map φ_α , the n numbers $v[\varphi_\alpha^i]$ are components of a vector $\hat{v} \in \mathbb{R}^n$. Now we can let $\hat{\gamma}(t) = \varphi_\alpha(p) + t\hat{v}$ and $\gamma(t) = \varphi_\alpha^{-1}(\hat{\gamma}(t))$. It is easy to show that this is a linear inverse to the above mapping. ■

Exercise 3.4. Check that the definition of a tangent vector and all of the following constructions are consistent, i.e. independent of the choices of charts φ_α or representatives $\gamma(t)$ of equivalence classes.

Definition 3.6.4 (Differential/push-forward). Let $F \in C^\infty(M, N)$ and $p \in M$. The push-forward of F at p is the linear map $D_p F = F_{*p} : T_p M \rightarrow T_{F(p)} N$ defined via

$$F_{*p}([\gamma]_p) = [F \circ \gamma]_{F(p)}, \quad (3.6.1)$$

or equivalently

$$F_{*p}(v)[f] = v[f \circ F] \quad (3.6.2)$$

for any $f \in C^\infty(N)$.

Remark 3.6.5 (Coordinate expressions for derivatives). In local coordinates (i.e. replacing F with some function $\hat{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$F_{*p}([\gamma]_p) = \frac{d}{dt} \Big|_{t=0} \hat{F}(\gamma(t)). \quad (3.6.3)$$

and

$$\hat{F}_{*p}(v)[f] = \frac{d}{dt} \Big|_{t=0} f(F(\gamma(t))). \quad (3.6.4)$$

Theorem 3.6.6. Consider smooth maps between manifolds $M \xrightarrow{f} N \xrightarrow{g} P$. Then

1. $f_{*p} : T_p M \rightarrow T_{f(p)} N$ is linear;
2. $(g \circ f)_{*p} = g_{*f(p)} \circ f_{*p}$ (chain rule);
3. $(\text{id}_M)_{*p} = \text{id}_{T_p M}$;
4. if f is a local diffeo at x then f_{*p} is invertible and $(f_{*p})^{-1} = (f^{-1})_{*p}$ (the converse is also true by InMT);
5. In local coordinates, f_{*p} is represented by the Jacobi matrix of the local representative $f_{\beta\alpha}$;
6. $\dim T_p M = \dim M$, even for points on ∂M ;
7. InMT doesn't hold at the boundary.

Proof. The first few statements are trivial and are left as an exercise. 6) A local chart map φ_α can be viewed as a local diffeo to an open subset of \mathbb{R}^n . Since $T_{\varphi_\alpha(p)}(\mathbb{R}^n) = \mathbb{R}^n$ (because we know that equivalence classes of curves at a point of \mathbb{R}^n are uniquely characterized by their tangent vectors), we have a linear map $(\varphi_\alpha)_{*p} : T_p M \rightarrow \mathbb{R}^n$, which is an isomorphism by 4). Therefore the dimensions coincide. Nothing really changes at the boundary.

7) Consider the inclusion map $i : \mathbb{R}^n_+ \hookrightarrow \mathbb{R}^n$. Clearly $i_{*0} = \text{id}$ is invertible, but i itself is not a local diffeo because $i(B^n \cap \mathbb{R}^n_+)$ is the image of an open set that is not open in \mathbb{R}^n . ■

All of this implies that $T : (M, p) \mapsto T_p M$ is a covariant functor $T : \text{Man}_\bullet^\infty \rightarrow \text{Vect}_\mathbb{R}$. It respects morphisms (smooth maps) $f : M \rightarrow N$ between pointed¹⁴ manifolds because the following diagram commutes:

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ T_p \downarrow & \curvearrowright & \downarrow T_{f(p)} \\ T_p M & \xrightarrow{T_p f} & T_{f(p)} N. \end{array} \quad (3.6.5)$$

As we will show later, the tangent space over $p \in M$ in a certain sense “varies smoothly with p ” and together all of the tangent spaces comprise a single object called the *tangent bundle*. A similar thing happens to the functors T_p .

3.7 Rank

We have defined rank in Def. 3.4.2. However, that definition is not practical because it's an existence statement rather than a constructive “measurement”.

¹⁴The requirement that these manifolds be pointed is a mere technicality for the sake of the definition of a functor. Given a map $f : M \rightarrow N$ and a point $f \in M$, we can always create the pairs (M, p) and $(N, f(p))$ and treat f as a pointed map.

Recall that for a matrix $A \in \text{Mat}(n \times m, \mathbb{R})$ its rank $\text{rank } A$ is defined as the maximal number of linearly independent columns (equivalently, rows) in it. More abstractly, the rank of a linear operator is the dimension of its image, $\text{rank } A = \dim(\text{im}(A))$. The following theorem establishes the equality between the rank of a smooth map and the rank of its differential.

Theorem 3.7.1 (Rank theorem). *For $f \in C^\infty(M, N)$, $\text{rank}_p f = \text{rank}(f_{*p})$.*

Proof. This is the **InMT** in disguise. Since this theorem is local, we can safely replace M and N by open $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$.

Assume $\text{rank}(f_{*p}) = r$. Introduce coordinates $(x^1, \dots, x^r, y^1, \dots, y^{m-r})$ on U and $(v^1, \dots, v^r, w^1, \dots, w^{n-r})$ on V . In these coordinates, we can write

$$f(x, y) = (F(x, y), G(x, y)). \quad (3.7.1)$$

Without loss of generality (WLOG), $p = 0$, $f(p) = 0$, and $F_{*(0,0)}$ is a nonsingular square matrix.

Define $\varphi : U \rightarrow \mathbb{R}^m$ by $\varphi(x, y) = (F(x, y), y)$. Its differential at $(0, 0)$ is nonsingular by assumption, so by the **InMT** we locally have an inverse and by definition of φ , the composition of φ^{-1} with f will have the form

$$f \circ \varphi^{-1} = (x, \tilde{G}(x, y)). \quad (3.7.2)$$

Finally, by a trivial change of coordinates $\psi(v, w) = (v, w - \tilde{G}(v, 0))$ we can bring the local form of f to

$$\psi \circ f \circ \varphi^{-1}(x, y) = (x, 0), \quad (3.7.3)$$

which means that $\text{rank}_p f = r$.

The converse direction is obvious. ■

Proposition 3.7.2. *For any smooth map $f \in C^\infty(M, N)$ the set of points p where the rank is maximal, i.e. $\text{rank}_p f = \min(\dim M, \dim N)$, is open in M .*

Proof. Let $\dim M = m$ and $\dim N = n$. The mapping $D : p \mapsto \hat{f}_{*p}$ is a continuous (in fact, smooth) map $U_\alpha \subset M \rightarrow \text{Mat}(n \times m, \mathbb{R})$. The subset of all matrices of maximal rank is open in $\text{Mat}(n \times m, \mathbb{R})$ (essentially because small perturbations can't make linearly independent vectors dependent¹⁵). By continuity of D , the set of p 's at which the rank is maximal is also open (in every U_α and thus in all of M). ■

¹⁵The converse is not true, which is why the set of matrices of any fixed non-maximal rank is not open.

Therefore we see that maps of maximal rank are of special significance. Moreover, they fall into two categories: if $\dim M \leq \dim N$, then the maximal rank equals $\dim M$; if $\dim M \geq \dim N$, the maximal rank is $\dim N$. In the first case, matrices of maximal rank are exactly the injective ones, whereas in the second case they are the surjective ones.

Definition 3.7.3 (Smooth immersions and submersions). Among all maps of *constant rank* we single out two special classes.

1. $f \in C^\infty(M, N)$ is called an immersion if $f_{*p} : T_p M \rightarrow T_{f(p)} N$ is injective for all p . Equivalently, f is an immersion iff it has everywhere maximal rank and $\dim M \leq \dim N$.
2. $f \in C^\infty(M, N)$ is called a submersion if $f_{*p} : T_p M \rightarrow T_{f(p)} N$ is surjective for all p . Equivalently, f is a submersion iff it has everywhere maximal rank and $\dim M \geq \dim N$.

Remark 3.7.4. As is clear from Theorem 3.7.1, a submersion is a map that is locally an orthogonal projection onto a subspace (in some coordinates), and an immersion is one that is locally an inclusion as a lower-dimensional subspace. A map that is both a submersion and an immersion is a local diffeomorphism. Note that “locality” here is understood both in domain and in range – we have to pick out open neighborhoods of both x and $f(x)$. An immersion for which this property holds globally on the entire domain (i.e. locally only in the target manifold) is just an injective immersion. Meanwhile, the attempt to define maps that are “projections locally in the target” will lead us to the concept of fiber bundles, and it will not be enough for them to be surjective submersions.

Theorem 3.7.5 (Global rank theorem). Let $f \in C^\infty(M, N)$ have constant rank r . Then

1. if f is injective, it is an immersion;
2. if f is surjective, it is a submersion;
3. if f is bijective, it is a diffeomorphism.

Proof.

1. Suppose it is not an immersion. Then $r < \dim M$, which means that in some coordinates we have $\hat{f}(0, \dots, 0, \epsilon) = (0, \dots, 0)$ for all sufficiently small ϵ . This contradicts injectivity.
2. Suppose it is not a submersion. Then $r < \dim N$. There exists a (countable) open covering $\{U_\alpha\}$ of M such that $f(U_\alpha)$ are ball-slices $B^n \cap \{y^{r+1} = y^{r+2} = \dots = y^n = 0\}$. We have $f(M) = \cup_\alpha f(U_\alpha)$, but at the same time $f(U_\alpha)$'s are *nowhere dense* (i.e. closure has empty interior). By Baire category theorem, a countable union of nowhere dense sets has an empty interior, which implies $\text{int } f(M) = \emptyset$. This contradicts surjectivity, because $\text{int } N \neq \emptyset$.
3. Obvious (a bijective local diffeo is a diffeo).

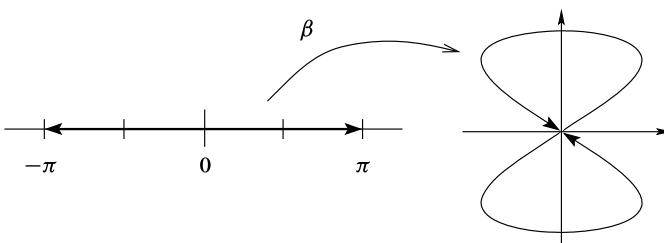


Fig. 3.1: The figure-eight curve $\beta : (-\pi, \pi) \rightarrow \mathbb{R}^2$ given by $\beta(t) = (\sin 2t, \sin t)$ is an immersion but not an embedding. Its domain is open but its image is closed.

3.8 Immersions and embeddings

As we described above, immersions are maps with injective differentials. Immersions themselves don't need to be injective, and even if they are, their images, as topological subspaces of the codomain, are not even always homeomorphic to the domain. For example, the map in Figure 3.1 is an immersion whose image is not homeomorphic to the domain. Maps that do respect subspace topology in this way are called embeddings.

In the topological category, the situation is different. A topological immersion is the intuitive analog of the smooth immersion: the injectivity of differentials is replaced by a local homeomorphism on the domain.

Definition 3.8.1 (Topological immersion). A continuous map $f : X \rightarrow Y$ is called a topological immersion if for every $x \in X$ there is an open neighborhood $U \subset X$ of x such that $f|_U \rightarrow f(U)$ is a homeomorphism, where $f(U)$ is taken to have the subspace topology inherited from Y .

A topological embedding then is a global version of a topological immersion: the entire map has to be a homeomorphism onto the image. In particular, it has to be injective.

Definition 3.8.2 (Topological embedding). A topological embedding is a continuous map $f : X \rightarrow Y$ that is a homeomorphism onto its image (where the image $f(X) \subset Y$ is taken to have the subspace topology in the target). In other words, it is an injective map such that the topology on the domain must coincide with the initial topology induced by the map.

In this way the concept of a topological embedding map is dual to that of a topological quotient maps. In the smooth category the picture is a bit different. First, the intuitive smooth version of a quotient map is a smooth submersion that is a topological quotient map. However, as we will see in Section 3.10, every smooth submersion is automatically a quotient map, so the second requirement can be dropped. On the other hand, smooth immersions are not

always topological embeddings (although they are topological immersions). This is why smooth submersions and smooth immersions are not dual concepts, but instead the concept dual to smooth submersions is the following concept of a smooth embedding.

Definition 3.8.3 (Smooth embedding). A smooth embedding is a smooth immersion that is topological embedding.

Example 3.8.4. 1. Much like the example in Figure 3.1, the mapping of \mathbb{S}^1 onto a figure eight in \mathbb{R}^2 can be represented by a parametric map $f(t) = (f_x(t), f_y(t))$. It is an immersion provided that $f'(t)$ doesn't vanish anywhere. However, it is not an embedding because the image (figure eight) is not homeomorphic to the circle.

2. Excluding the one point on the circle that is mapped to the center point of the figure eight in the last example, we obtain a map $f : (0, 1) \rightarrow \mathbb{R}^2$ whose image is still the whole figure eight. This is an injective immersion. However, it is still not an embedding because the image is compact in the subspace topology, whereas the open interval is not a compact space.

3. $\gamma : \mathbb{R} \rightarrow \mathbb{T}^2$ given by $\gamma(t) = (e^{2\pi i t}, e^{2\pi i \alpha t})$ for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is an injective immersion. However, its image is dense in the torus, which means that its subspace topology is very different from that of the real line. Namely, there exist convergent sequences of points in the image whose pre-images don't converge on the real line. This means that γ^{-1} is not continuous and γ is not an embedding.

4. The map $f(x) = (0, x^3)$ is a topological embedding $\mathbb{R} \hookrightarrow \mathbb{R}^2$ but not a smooth embedding because its differential has rank zero at $x = 0$. Essentially this restriction insures that smooth structures induced via smooth embeddings coincide with the original ones. The smooth structure induced via $(0, x^3)$ from \mathbb{R}^2 (with the standard structure) would give one of the non-standard smooth structures on \mathbb{R} .

Proposition 3.8.5. An injective immersion $f \in C^\infty(M, N)$ is a smooth embedding if any of the following hold:

1. f is open or closed;
2. f is proper;
3. M is compact;
4. $\partial M = \emptyset$ and $\dim M = \dim N$.

Proof. See [7, Prop 4.22]. ■

Theorem 3.8.6 (Local embedding theorem). $f \in C^\infty(M, N)$ is an immersion iff it is a local embedding (i.e. every point in M has a neighborhood on which the restriction of f becomes an embedding).

Proof. See [7, Thm 4.25]. ■

3.9 Submanifolds

Immersions and embeddings give rise to the notions of immersed and embedded submanifolds.

Definition 3.9.1. Let S and M be smooth manifolds and let $i : S \rightarrow M$ be a continuous map.

This triple $S \xrightarrow{i} M$ is called:

1. an immersed submanifold if i is a smooth immersion. We write $S < M$.
2. a weakly embedded submanifold, or *initial*, if it is immersed and every smooth map $f : N \rightarrow M$ such that $\text{im } f \subset S$ restricts in the codomain to a smooth map $f : N \rightarrow S$.
3. an embedded submanifold if i is a smooth embedding map. Often $S \subset M$ with the subspace topology. Then i is required to be a only a topological embedding (and a smooth structure on S can be induced via i). We write $S \triangleleft M$.
4. a properly embedded submanifold if i is a proper smooth embedding map. We write $S \leqslant M$.

The codimension of a submanifold is defined as $\text{codim } S = \dim M - \dim S$.

Note that sometimes immersed “submanifolds” are defined as images of arbitrary immersions, but we avoid that definition since such subsets generally don’t admit any manifold structure (since they can self-intersect).

- Example 3.9.2.**
1. Any open subset is a codimension zero embedded submanifold.
 2. If $F : M \rightarrow N$ is an embedding, then $F(M) \triangleleft N$ in the subset sense. i.e. embedded submanifolds inside N are exactly the images of embedding maps.
 3. If $S \triangleleft M$ as a subset, then S is closed iff $S \leqslant M$.
 4. If $S \triangleleft M$ and S is compact, then $S \leqslant M$.
 5. The figure eight immersion $t \mapsto (\sin 2t, \sin t)$ is a weak embedding but not an embedding.
 6. If $f \in C^\infty(N, M)$, then its graph $\text{graph}(f) \subset N \times M$ is an embedded submanifold.

Proposition 3.9.3 ([7, Prop. 5.18]). *If $F : M \rightarrow N$ is an injective smooth immersion, then $F(M)$ has a unique topology and smooth structure such that it is a smooth submanifold of N and such that $F : M \rightarrow F(M)$ is a diffeomorphism onto its image.*

Proof. The topology and smooth structure on $F(M)$ are induced by F itself (i.e. $U \subset N$ is open iff $F^{-1}(U) \subset M$ is open, and the charts on $F(M)$ are $(F(U), \varphi \circ F^{-1})$ for every chart (U, φ) on M), and they obviously turn F into a diffeomorphism onto its image, and these are the only topology and smooth structure with this property. The inclusion $F(M) \hookrightarrow N$ can be written as a composition $F(M) \xrightarrow{F^{-1}} M \xrightarrow{F} N$. As a composition of a diffeomorphism and a smooth immersion, it is a smooth immersion. ■

Proposition 3.9.4. A subset $S \subset M$ is an embedded submanifold iff every point $p \in S$ has a local “slice chart”, that is, a local chart map φ on a neighborhood $U_p \subset M$ such that its image is a slice ball: $\varphi(S \cap U_p) = \{(x^1, \dots, x^k, 0, \dots, 0)\}$, where k is a fixed number called the dimension of S . In particular, given the existence of slice charts, there is a unique smooth structure on S (given by a codomain restriction of the slice atlas) compatible with the subset topology such that $S \triangleleft M$.

Proof. It suffices to check that slice charts, whose existence was proved above, form a smooth atlas on S . This also proves that, given a topologically embedded submanifold S in a smooth manifold M , one can induce a smooth structure on S , essentially by defining the slice atlas. See all details in [7, Thm 5.8]. ■

Corollary 3.9.5. ∂M has slice charts by definition, therefore it is an embedded submanifold.

Proposition 3.9.6. An embedded submanifold $S \subset M$ is properly embedded iff it is a closed subset of M .

Proof. Exercise. ■

Theorem 3.9.7 (Level Set Theorem). If $f \in C^\infty(M, N)$ has constant rank r then the level set $f^{-1}(q)$ for any $q \in N$ is a properly embedded codimension r submanifold.

Proof. By definition of rank, there are coordinates in which the local representative is $\hat{f}(x^1, \dots, x^m) = (x^1, \dots, x^r, 0, \dots, 0)$. Therefore, assuming q is the origin of its coordinate system, $f^{-1}(q) = (0, \dots, 0, \underbrace{y^1, \dots, y^{m-r}}_{\text{arbitrary}})$. This means that $f^{-1}(q)$ is a local ball slice and

therefore an embedded submanifold. It is closed as a pre-image of a closed set under a continuous map, and therefore properly embedded. ■

Corollary 3.9.8. Each level set of a smooth submersion $f : M \rightarrow N$ is properly embedded and of dimension $\dim N$.

Corollary 3.9.9 (Regular level set theorem). Every regular level set (i.e. level set of a regular value) of a smooth map $f : M \rightarrow N$ is a properly embedded submanifold of codimension $\dim N$.

Proof. Since the set of all regular points, being the set of points where f_* has maximal rank, is open (Proposition 3.7.2), the level set is contained in it together with an open neighborhood U of itself: $f^{-1}(q) \subset U \subset M$. Then $f|_U$ is a smooth submersion and we can use the last Corollary. ■

Proposition 3.9.10. 1. All embedded submanifolds $S \subset M$ have local defining functions, i.e. $\varphi : C^\infty((U \subset M, \mathbb{R})$ such that $\varphi^{-1}(0) = S \cap U$.
 2. An immersion is a local embedding (i.e. its restriction to a sufficiently small neighborhood $U \subset S$ of a point $p \in S$ is an embedding¹⁶).

Proof. 1. Take a local slice chart such that $S|_U = \{(x^1, \dots, x^k, 0, \dots, 0)\}$ and define $\varphi = (x^{k+1})^2 + \dots + (x^m)^2$.
 2. By Theorem 3.8.6. ■

Theorem 3.9.11 (Extensions of functions). 1. If $S \lhd M$ as a subset, then any $f \in C^\infty(S)$ can be extended to a $\tilde{f} \in C^\infty(U)$ for some open neighborhood $S \subset U \subset M$ (possibly $U = S$).
 2. If in the above, $S \trianglelefteq M$, then there is extension to the whole M .

Proof. 1. These extensions are possible locally because they are possible in \mathbb{R}^n (extensions off a coordinate ball slice; no extension needed if the dimension of the slice is n). Then we use a POU to glue them together around S much like in Theorem 3.3.4.
 2. In this case we can just refer to the extension lemma 3.3.4 for closed subsets since any properly embedded manifold is a closed subset in M . ■

Exercise 3.5. Show that the two parts of the last theorem are in fact sufficient, i.e. they are criteria for S being embedded or properly embedded, respectively.

Theorem 3.9.12 (Restrictions of functions [7, Thms. 5.27-28]). Let $S \subset M$ as a subset and $f \in C^\infty(M, N)$. Then:

1. the restriction $f|_S : S \rightarrow N$ is smooth;
2. if $K \subset N$, $f(M) \subset K$, and f is continuous as a map $M \rightarrow K$, then the codomain restriction $f|^K : S \rightarrow K$ is also smooth.
3. if $K \lhd N$ and $f(S) \subset K$, then the double restriction $f|_S^K : S \rightarrow K$ is also smooth.

Proof. The idea is to show that $f|_S = f \circ i$, where $i : S \hookrightarrow M$ is the smooth inclusion map. This is smooth by virtue of being a composition of smooth maps.

For the rest, see the proof in [7, Thms. 5.27-30]. ■

Note that to be able to restrict a smooth map in range to an immersed manifold S , we need it to be continuous as a map to S with respect to the subspace topology on S . Immersed

submanifolds which allow for restrictions of smooth maps in range without any such restrictions are exactly the weakly embedded ones.

Remark 3.9.13. The terminology around submanifolds is quite confusing and generally only the concept of embedded submanifolds is consistent across different authors. Our use of the term ‘immersed submanifold’ (meaning an image of a smooth injective immersion) is consistent with [7]. As indicated in Theorem 3.9.12, a closely related concept is that of weakly embedded (as in [7]), or initial (as in [16]), submanifolds. The definition of weakly embedded submanifolds can be thought of as a universal property, and as we’ve shown above, every embedded submanifold is weakly embedded, however the converse is not true. In other words, we have the following strict inclusions:

$$\text{Proper embeddings} \subset \text{Embeddings} \subset \text{Weak embeddings} \subset \text{Immersions}. \quad (3.9.1)$$

Theorem 3.9.14 ([7, Thm. 5.31]). *If $S \lhd M$ is embedded, then the subspace topology on S and the smooth structure defined by slice charts are the only topology and smooth structure with respect to which S is an embedded or even immersed manifold.*

Theorem 3.9.15 ([7, Thm. 5.32]). *If $S < M$ is immersed, then for a given topology on S there is only one smooth structure making S into an immersed submanifold.*

It is certainly possible for a given subset of M to have more than one topology making it into an immersed submanifold. However, for weakly embedded submanifolds we have a stronger uniqueness result.

Theorem 3.9.16 ([7, Thm. 5.33]). *If $S \subset M$ is a weakly embedded submanifold, then S has only one topology and smooth structure with respect to which it is an immersed submanifold.*

- Exercise 3.6.**
1. Show $\partial f^{-1}(q) = f^{-1}(q) \cap \partial M$ for any $f \in C^\infty(M, N)$ and $q \in N$ that is a regular value for both f and $f|_{\partial M}$.
 2. Give an example of an immersion $S < M$ and a function $f \in C^\infty(M)$ such that the restriction $f|_S$ is not smooth.

3.10 Submersions

Recall that a split epimorphism is an epi that has a “right inverse”, or a section/coretraction. In the category Man^∞ , epimorphisms are exactly the smooth maps whose images are dense in the target. Split epimorphisms, however, have to be truly surjective.

Definition 3.10.1 (Local section). A local section of a continuous (or smooth) map $\pi : M \rightarrow N$ is a continuous (resp. smooth) map $\sigma : U \rightarrow M$ defined on an open set $U \subset N$ such that $\pi \circ \sigma = \text{id}_U$.

This allows us to introduce a topological version of submersions. It looks drastically different from the definition in the smooth case.

Definition 3.10.2 (Topological submersion). A continuous map $f : X \rightarrow Y$ is called a topological submersion if for every $x \in X$ there exists an open neighborhood $V \subset Y$ of $f(x)$ and a local section $\sigma_V : V \rightarrow X$ of f that contains x in its image, i.e. $x \in \sigma(V)$ and $f \circ \sigma = \text{id}_V$.

An alternative way to define topological submersions would be to read the Rank Theorem 3.7.1 (also known as the Local Submersion Theorem in the special case of submersions) combined with our definition of rank (Def. 3.4.2) to write down a topological characterization of submersions: around every point $x \in X$ there is a neighborhood $U \subset X$ of x and a neighborhood $V \subset Y$ of $f(x)$ and a homeomorphism $\phi : U \rightarrow V \times Z$ with some space Z such that f is locally equivalent to the projection onto the first component of $V \times Z$, i.e. $f|_U = \text{pr}_1 \circ \phi$. This definition turns out to be a special case of the one given above.

We will now show that smooth submersions are the exact analog of topological submersions: a smooth map is a submersion iff each point in the domain admits a smooth local section through it.

Theorem 3.10.3 (Local Section Theorem [7, Thm. 4.26]). $\pi \in C^\infty(M, N)$ is a submersion iff every $p \in M$ lies in the image of a smooth local section of π .

Proof. For the forward direction, we have in some local coordinates $f_{\alpha\beta}(x^1, \dots, x^m) = (x^1, \dots, x^n)$, where we assume that p is the origin of the coordinate system. Therefore the map locally defined by $\sigma_{\alpha\beta}(y^1, \dots, y^n) = (y^1, \dots, y^n, 0, \dots, 0)$ is indeed a local section and its image passes through p .

For the backward direction, let $p \in M$ and find a local section σ such that $p \in \sigma(U)$. Then $p = \sigma(q)$ for some $q \in N$ and $\pi \circ \sigma = \text{id}_U$. Therefore, by taking the differential, $\pi_* p \sigma_{*q} = \text{id}_{T_q N}$, which implies that π_* is surjective. ■

Next we would like to find the smooth analog for the concept of a quotient map. Recall that a *topological quotient map* is a surjective, strongly continuous map (i.e. V is open iff $f^{-1}(V)$ is open). Openness implies strong continuity, thus open surjective maps are necessarily quotient maps, but not every topological quotient map is open (e.g. the gluing $\mathbb{D}^n / \partial \mathbb{D}^n$ of the boundary of a ball that gives a sphere is not open). In fact, a surjective continuous map q is a quotient map iff it takes open sets that are unions of fibers of q to open sets. However, these subtleties turns out to be irrelevant in the smooth case as the natural smooth analogs of surjective continuous maps, *submersions*, turn out to be open by the following proposition.

Proposition 3.10.4. Any smooth surjective submersion is an open topological quotient map.

Proof. Call the map in question π . Let $W \subset M$ be an open set and $p \in W$. Pick a local section σ passing through p , and pick a point $y \in \sigma^{-1}(W)$. Then $y = \pi \circ \sigma(y) \in \pi(W)$. Therefore $\sigma^{-1}(W)$ is an open neighborhood of $\pi(p)$ contained inside $\pi(W)$. Since $\pi(p)$ can be any point in $\pi(W)$, this means that $\pi(W)$ is open, making π an open map.

We have thus proved that a smooth submersion is open. A surjective submersion is then a topological quotient map. ■

This result motivates the definition of quotient maps in the smooth category Man^∞ . Namely, we only need to require the submersion property (i.e. “infinitesimal surjectivity”) on top of simple surjectivity. By the above result, there is no need to require them to be topological quotient maps.

Definition 3.10.5 (Smooth quotient map). A *smooth* quotient map is a surjective submersion.

We have shown that smooth submersions are always open (unlike topological ones). Thus, given a smooth submersion $f : M \rightarrow N$, we know that $f(M)$ is open in N and thus an embedded submanifold. Hence we can safely restrict f in its range and obtain a smooth quotient map. With this remark, *smooth submersions and smooth quotient maps are functionally the same thing*.

Remark 3.10.6. $\pi \in C^\infty(M, N)$ is a smooth quotient map iff it is a topological quotient map and the given topology and smooth structure on N are the unique ones that make π into a smooth submersion. See [7, Problem 4-7].

What we have learned can be summarized in the following table, with the lower right cell being a teaser of what’s to come.

	<i>Injective version</i>	<i>Surjective version</i>
<i>Basic version</i>	Immersion	Submersion
<i>Stronger version</i>	Injective Immersion	Surjective Submersion (a.k.a. Quotient Map)
<i>Strongest version</i>	Embedding	Fiber Bundle

This table also indicates two possible directions a course in differential geometry can take. The theory of embedding maps (including Whitney’s and Nash’s embedding theorems, among many others) is fairly sophisticated, and so is the theory of fiber bundles. However, we will mostly follow the “submersion route” due to its direct physical relevance.

4 Covering spaces

4.1 Homotopy lifting

A very special type of a submersion is a covering map (which, as we will see, is a special case of a fiber bundle). It turns out that the smooth structure doesn't add anything to the study of these objects, so we will work in the category of topological spaces. The content of this section, with minor additions, is borrowed from the famous book by Bredon [19], however we can also recommend [8, Ch. 11] for a quicker treatment of the case of topological manifolds.

Definition 4.1.1 (Topological covering map). A continuous map $E \xrightarrow{\pi} M$, where E is called the *total space*, M the *base space* and π the projection, is a *topological covering map* if it is surjective and every $m \in M$ has an open neighborhood U such that $\pi^{-1}(U)$ consists of at most countably many connected components \tilde{U}_i and each restriction $\pi|_{\tilde{U}_i} : \tilde{U}_i \rightarrow U$ is a homeomorphism. Such U_i are called *elementary* or *evenly covered*, and we say that π is *trivial* over U_i . The pre-image $\pi^{-1}(m)$ of a point in the base is called the *fiber above* m .

Definition 4.1.2 (Smooth covering map). A smooth map $E \xrightarrow{\pi} M$ between smooth manifolds E and M is a *smooth* covering map if it is a topological covering map and each restriction $\pi|_{\tilde{U}_i}$ is a diffeomorphism.

In other words, the total space E is locally isomorphic to parts of M , but the pre-image of an open set in M can consist of countably many copies of itself embedded in E .

Proposition 4.1.3. *The following properties hold for any smooth covering map π .*

1. *The neighborhood U whose existence is required in the definition has to be connected.*
2. *Covering maps are local diffeomorphisms, submersions, open maps, and quotient maps.*
3. *If π is injective then it is diffeo.*
4. *For $p \in \pi^{-1}(q)$ and U a neighborhood of q as in the definition, there is a unique local section $\sigma : U \rightarrow E$ of π such that $\sigma(q) = p$.*

Proof.

1. Obvious because every connected component of the pre-image can't be diffeomorphic to a disconnected U .
2. Obviously a local diffeo. Submersion because local diffeo. Open for the same reason (or because a submersion). Quotient because surjective and submersion.
3. If π is injective, it's an immersion by the Global Rank Theorem 3.7.5. A map that is an immersion and a submersion is a diffeo.
4. A section is guaranteed by definition. Just let \tilde{U} be the connected component of $\pi^{-1}(U)$ that contains p . Then $\pi|_{\tilde{U}}$ is a diffeomorphism whose inverse is obviously a section passing through p defined on U . Now, suppose there is another section $\sigma' : U \rightarrow E$

with the same properties. Then $\sigma'(U) \subset \tilde{U}$ since U is connected and $\sigma'(q) = p$. But then both σ and σ' are right inverses to the bijective map $\pi|_{\tilde{U}}$, so they have to coincide. ■

For now we consider arbitrary covering spaces, not just smooth ones.

Lemma 4.1.4 (Lebesgue Lemma). *Let X be a compact metric space and let $\{U_\alpha\}_\alpha$ be an open covering of X . Then there exists a constant $\delta > 0$ such that any subset $A \subset X$ of diameter $\text{diam}(A) < \delta$ is contained in U_α for some α .*

Proof. For each $x \in X$ there is an $r(x) > 0$ such that the ball $B_{2r(x)}(x)$ is contained inside U_α for some α . Then by compactness X is covered by a finite number of the balls $B_{r(x)}(x)$, say for $x = x_1, \dots, x_n$. Define $\delta = \min\{r(x_i) \mid i = 1, \dots, n\}$. Now if $\text{diam}(A) < \delta$, then there is an i such that $\text{dist}(A, x_i) < r(x_i)$. Since $\text{diam}(A) < \delta \leq r(x_i)$, by the triangle inequality $\text{dist}(a, x_i) < 2r(x_i)$. Thus $A \subset B_{2r(x_i)}(x_i) \subset U_\alpha$ for some α . ■

Lemma 4.1.5 ([19, Lem. 3.2]). *Let W be a topological space and let $\{U_\alpha\}_\alpha$ be an open covering of $W \times I$ (where $I = [0, 1]$). Then for any $w \in W$ there is a neighborhood $N \subset W$ of w and a positive integer n such that $N \times [i/n, (i+1)/n] \subset U_\alpha$ for some α , for each $0 \leq i < n$.*

Proof. $\{w\} \times I$ can be covered by a finite refinement of $\{U_\alpha\}_\alpha$ of the form $\{N_i \times V_i\}_i$ by compactness of I . The Lebesgue Lemma 4.1.4 implies the existence of n such that $[i/n, (i+1)/n]$ is contained in one of the V_j . Now take $N = \bigcap_i N_i$. ■

Theorem 4.1.6 (Path Lifting Property). *Consider a topological covering space $E \xrightarrow{\pi} M$ and let $\gamma : [0, 1] \rightarrow M$ be a path in M . Given a point p in the fiber above $\gamma(0)$, there is a unique path $\tilde{\gamma} : [0, 1] \rightarrow E$ such that $\tilde{\gamma}(0) = p$ and $\pi \circ \tilde{\gamma} = \gamma$. We call $\tilde{\gamma}$ the lift of γ based at p .*

Proof. By the Lebesgue Lemma, there is an n such that each $\gamma([i/n, (i+1)/n])$ lies in an elementary set. Then we can lift by induction in i on each elementary set using the local homeomorphism. At each stage of the induction, the lift is already defined at the left-hand endpoint, leading to uniqueness since it singles out a connected component in the pre-image of the elementary set. ■

This property is a characteristic property of covering maps and is key in many applications of this theory, such as complex analysis, where the Riemann surfaces of analytic functions are obtained via analytic continuation along paths in the complex plane. The uniqueness of

analytic continuation is what guarantees the existence of such a covering map. In fact, the lifting property can be used to define the notion of a covering map.

Now we show the more general form of the lifting property.

Theorem 4.1.7 (Covering Homotopy Property/Homotopy Lifting Property). *Given a covering space $\pi : E \rightarrow M$, a locally connected space W , a homotopy $F : W \times I \rightarrow M$, and a map $\tilde{f}_0 : W \rightarrow E$ lifting $f_0 = F(\cdot, 0)$, there exists a unique homotopy $\tilde{F}_t : W \times I \rightarrow E$ of \tilde{f}_0 that lifts F .*

Proof. Define \tilde{F} on each $\{w\} \times I$ as the unique lifting of a path from Theorem 4.1.6. Now we need to show continuity in w . By Lemma 4.1.5 we can find a connected neighborhood N of w in W and an integer n so that $F(N \times [i/n, (i+1)/n])$ is in some elementary set U_i . Assuming that \tilde{F} is continuous on $N \times \{i/n\}$ we see that $\tilde{F}(N \times \{i/n\})$, being connected, must be contained in a single connected component, say V of $\pi^{-1}(U_i)$. But then on $N \times [i/n, (i+1)/n]$, the lifting \tilde{F} must be F composed with the inverse of the homeomorphism $\pi|_V : V \rightarrow U_i$ (again using connectivity). But this means that \tilde{F} is continuous on all of $N \times [i/n, (i+1)/n]$. A finite induction then shows that \tilde{F} is continuous on each $N \times I$, and hence everywhere. ■

Remark 4.1.8. With a slight improvement of the proof, the condition of local connectedness for W can be dropped.

The uniqueness part of this property is special to covering spaces. In topology, *fibrations* $\pi : P \rightarrow M$ are generally defined as maps that have the homotopy lifting property (but without uniqueness).

Definition 4.1.9 (Relative homotopy). Let X, Y be topological spaces and $K \subset X$ a subspace of X . Two maps $f_0, f_1 : X \rightarrow Y$ are said to be homotopic *relative* to A if there exists a homotopy $F : X \times I \rightarrow Y$ such that $F(a, t) = f_0(a) = f_1(a)$ is constant along I for each $a \in A$. We say that this is a homotopy “rel A ”. In particular, if A is a point, this is pointed homotopy that we used in defining homotopy groups. If $A = \{0, 1\} \subset I$, this introduces the notion of *homotopy of paths with fixed endpoints*.

Corollary 4.1.10 ([19, Cor. 3.7]). *If $\gamma_0 \sim \gamma_1$ are two paths homotopic in M and $\tilde{\gamma}_0, \tilde{\gamma}_1$ are their respective lifts in the covering space $\pi : E \rightarrow M$, then their endpoints coincide, $\tilde{\gamma}_0(1) = \tilde{\gamma}_1(1)$, and $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ are homotopic with fixed endpoints.*

Corollary 4.1.11. *Let $\pi : E \rightarrow M$ be a covering map. If γ is a trivial loop (i.e. homotopic to a constant loop) in M then any lifting of γ to a path in E is also a trivial loop.*

All of these properties of covering maps with respect to homotopies allow us to pass instead to equivalence classes of homotopic paths, which will transform the study of covering spaces

as topological objects into the study of their algebraic functors, namely the fundamental group.

Corollary 4.1.12 ([19, Cor. 3.7]). *Let $\pi : E \rightarrow M$ be a covering map and $\pi(p_0) = x_0$. Then*

$$\pi_* : \pi_1(E, p_0) \rightarrow \pi_1(M, x_0) \quad (4.1.1)$$

is a monomorphism whose image consists of the classes of loops at x_0 in M which lift to loops at p_0 in E .

Corollary 4.1.13 (Monodromy Theorem). *If γ is a loop in M based at x_0 which lifts to a loop in E based at p_0 then any loop homotopic to γ with fixed endpoints also lifts to a loop in E based at p_0 . That is, lifting to a loop is a property of the class $[\gamma] \in \pi_1(M, x_0)$.*

Corollary 4.1.14 ([19, Cor. 3.9]). *If a Hausdorff, path-connected, and locally path-connected space M has a nontrivial covering space (i.e. one which is not just a direct product of M with a discrete space) then $\pi_1(M, x_0) \neq 1$.*

Proof. Let p_0, p_1 be two points in the fiber above x_0 and let γ be a path between them. The $\pi \circ \gamma$ is a loop in M which does not lift to a loop in E . Then by Corollary 4.1.12 it follows that $[\pi \circ \gamma] \in \pi_1(M, x_0)$ is not in the image of $\pi_1(E, p_0)$ and hence it is a nontrivial element. ■

As a consequence of Corollary 4.1.14 we now know several spaces having nontrivial fundamental groups: the circle, the Klein bottle, and the projective plane. Let us start with the circle.

Definition 4.1.15 (Degree on a circle). Consider the covering space of \mathbb{S}^1 defined by the exponential map $p : \mathbb{R} \rightarrow \mathbb{S}^1$, $p(t) = e^{2\pi i t}$. Let $\gamma : I \rightarrow \mathbb{S}^1$ be a loop based at $1 \in \mathbb{S}^1$. Let $\tilde{\gamma} : I \rightarrow \mathbb{R}$ be a lifting of γ such that $\tilde{\gamma}(0) = 0$. Then $\tilde{\gamma}(1) = n \in p^{-1}(\{1\}) = \mathbb{Z}$. By Corollary 4.1.10, n depends only on the homotopy class $[\gamma] \in \pi_1(\mathbb{S}^1)$. This integer n is called the degree of γ and we write $n = \deg(\gamma)$.

Theorem 4.1.16 (Fundamental group of the circle). $\deg : \pi_1(\mathbb{S}^1) \rightarrow \mathbb{Z}$ is an isomorphism.

Proof. First we show that \deg is a homomorphism. Given two loops $\beta, \gamma : I \rightarrow \mathbb{S}^1$ and their lifts $\tilde{\beta}, \tilde{\gamma}$ both starting at $0 \in \mathbb{R}$, we have $\tilde{\beta}(1) = \deg(\beta) = n$ and $\tilde{\gamma}(1) = \deg(\gamma) = m$. Define $\tilde{\gamma}'(t) = \tilde{\gamma}(t) + n$. Then $\tilde{\gamma}'(0) = \tilde{\beta}(1)$, so $\tilde{\beta} \bullet \tilde{\gamma}'$ is defined, covers $\beta \bullet \gamma$, and $\tilde{\beta} \bullet \tilde{\gamma}'(1) = \tilde{\gamma}'(1) = \tilde{\gamma}(1) + n = m + n$, as claimed.

Second, \deg is surjective since a path from 0 to n in \mathbb{R} maps to a loop in \mathbb{S}^1 of degree n by definition.

Third, we saw that \deg is a monomorphism by showing its kernel is zero. Suppose $\gamma : I \rightarrow \mathbb{S}^1$ has degree 0. Then, for a lifting $\tilde{\gamma}$ we have $\tilde{\gamma}(1) = 0 = \tilde{\gamma}(0)$ so that $\tilde{\gamma}$ is a loop and represents an element $\tilde{\gamma} \in \pi_1(\mathbb{R}, 0) = 1$, since \mathbb{R} is contractible. Thus $[\gamma] = p_*([\tilde{\gamma}]) = p_*(1) = 1$. ■

Remark 4.1.17. The Path Lifting Property is our first example of a *connection*. A connection is a prescription for lifting curves from the base M into the total space. Alternatively, it is a prescription for lifting vector fields in the base to what is called *horizontal* vector fields in the total space. On covering spaces, as we see, the connection is unique due to the fact that the total space is locally diffeomorphic to the base.

The monodromy (how points in the fiber get permuted after a transport over a curve in the base) is also known as *holonomy* in the context of connections.

$$\begin{array}{ccc} & E & \\ \nearrow \tilde{f} & \curvearrowright & \downarrow \pi \\ W & \xrightarrow{f} & M \end{array} \quad (4.1.2)$$

Aside from homotopies, one can try to lift arbitrary mappings into the base space of a fibration. In general the existence of these liftings is difficult to establish. In the case of covering spaces, however, the answer is quite simple: a map from a space W into M can be lifted into the total space of $\pi : E \rightarrow M$ iff the images of all nontrivial loops in W are among those loops in M that can be lifted to E , and the lifting is unique. This is formalized in the following Theorem.

Theorem 4.1.18 (Lifting Theorem). *Let $\pi : E \rightarrow M$ be a covering map with $\pi(p_0) = x_0$. Assume that W is path-connected and locally path-connected and that $f : W \rightarrow M$ is a map with $f(w_0) = x_0$. Then a pointed map $\tilde{f} : (W, w_0) \rightarrow (M, x_0)$ such that $\pi \circ \tilde{f} = f$ exists iff $f_*(\pi_1(W, w_0)) \subset \pi_*(\pi_1(E, p_0))$. Moreover, \tilde{f} is unique.*

Proof. Let us construct \tilde{f} explicitly. Given $w \in W$, let $\lambda : I \rightarrow W$ be a path from w_0 to w . Then $f \circ \lambda$ is a path in M . Lift this to a path $\tilde{\lambda} : (I, 0) \rightarrow (E, x_0)$ and let $\tilde{f}(w) = \tilde{\lambda}(1)$. Then $\pi \circ \tilde{f}(w) = \pi(\tilde{\lambda}(1)) = f(\lambda(1)) = f(w)$.

To see that \tilde{f} is well defined, suppose λ' is another path in W from w_0 to w and let $\eta = (\lambda')^{-1}$ be the reverse path from w to w_0 . Then $\lambda \bullet \eta$ is a loop based at w_0 , and $(f \circ \lambda) \bullet (f \circ \eta)$ is a loop in M based at x_0 . Since

$$[f \circ \lambda \bullet f \circ \eta] = f_*([\lambda \bullet \eta]) \in \text{im}(f_*) \subset \text{im}(\pi_*), \quad (4.1.3)$$

$f \circ \lambda \bullet f \circ \eta$ lifts to a loop in E based at p_0 . The reverse of the portion of this lifting corresponding to η then is a lifting $\tilde{\lambda}'$ of λ' and $\tilde{\lambda}'(1) = \tilde{\lambda}(1)$, as required.

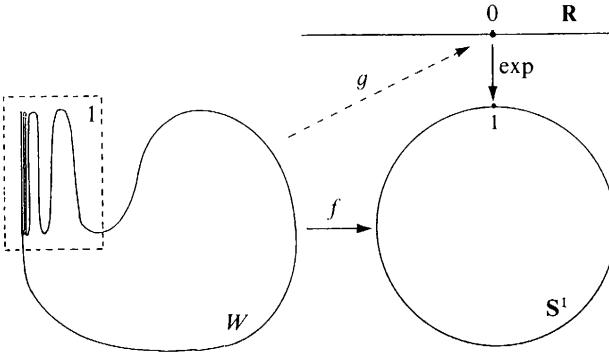


Fig. 4.1: The map f here collapses the entire “ $\sin 1/x$ ” part of the space on the left into the basepoint 1 on the circle. If a lifting map g into the real line existed, then the vertical part of the “ $\sin 1/x$ ” curve would have to be mapped to 0, whereas the wiggly part would map to 1 under g , making it discontinuous. From [19, Fig. III-6].

Next, we have to show that \tilde{f} is continuous. Let $w \in W$ and put $x = f(w)$. Let $U \subset \subset$ be an elementary neighborhood of x , and by local path-connectedness of W we can let V be a path-connected neighborhood of w such that $f(V) \subset U$. For any point $w' \in V$ we can construct a path from w_0 to w' by concatenating a given path λ to w with a path σ in V from w to w' . Since $f(V)$ is contained in an elementary set, the lifting of $f \circ \sigma$ is simply $f \circ \sigma$ composed with the local inverse of π that takes U to the component of $\pi^{-1}(U)$ that contains $\tilde{f}(w)$. This same component is used for all $w' \in V$ and it follows that \tilde{f} is continuous at w .

The converse direction of the theorem follows immediately from $f_* = \pi_* \circ \tilde{f}_*$. ■

Remark 4.1.19. Unlike in the homotopy lifting property (Theorem 4.1.7), the condition of local path connectedness for W cannot be dropped here. See an example in Figure 4.1.

Corollary 4.1.20 ([19, Cor. 4.2]). *Under the conditions of the Lifting Theorem 4.1.18, if W is also simply connected, then the lifting \tilde{f} always exists and is unique given a basepoint $p_0 \in E$.*

Corollary 4.1.21. *Given a covering space $E \xrightarrow{\pi} M$, there is a natural isomorphism of higher homotopy groups*

$$\pi_n(E, p_0) \cong \pi_n(M, x_0), \quad n \geq 2. \quad (4.1.4)$$

Proof. Injectivity follows from the uniqueness of the liftings of maps $\mathbb{S}^n \rightarrow M$ into E .

Surjectivity follows from the Lifting Theorem 4.1.18 combined with the fact that $\pi_1(\mathbb{S}^n)$ is trivial for $n \geq 2$ (Corollary 2.5.20). ■

Corollary 4.1.22. $\pi_n(\mathbb{S}^1) = 1$ for $n > 1$.

Proof. Any map $f : \mathbb{S}^n \rightarrow \mathbb{S}^1$ lifts to $\tilde{f} : \mathbb{S}^n \rightarrow \mathbb{R}$ by Corollary 4.1.20. But \tilde{f} is homotopically trivial since \mathbb{R} is contractible, so $f = \pi \circ \tilde{f}$ is also homotopically trivial. ■

4.2 Morphisms of covering spaces

Definition 4.2.1 (Category of covering spaces). The category Cov_M of covering spaces of a given base M (say, topological covering spaces) consists of all covering spaces of M . The morphisms between two covering spaces π and π' are morphisms $f : E \rightarrow E'$ that “map fibers to fibers”, i.e. $\pi' \circ f = \pi$, or equivalently if the following triangle commutes:

$$\begin{array}{ccc} E & \xrightarrow{f} & E' \\ \pi \searrow & \curvearrowright & \swarrow \pi' \\ M & & \end{array} \quad (4.2.1)$$

Lemma 4.2.2 ([19, Lem. 4.4]). Let W be connected and E be Hausdorff. Let $\pi : E \rightarrow M$ be a covering map and $f : W \rightarrow M$ a map. Let \tilde{f}_1, \tilde{f}_2 be two liftings of f into E . If $\tilde{f}_1(w) = \tilde{f}_2(w)$ for some point $w \in W$, then $\tilde{f}_1 \equiv \tilde{f}_2$.

Proof. Let $\tilde{f}_1(w) = \tilde{f}_2(w) = p$. Let U be an elementary open neighborhood of $f(w)$ in M . Let V be the component of $\pi^{-1}(U)$ containing p . Then $A = \tilde{f}_1^{-1}(V) \cap \tilde{f}_2^{-1}(V)$ is an open set in W and for $a \in A$ we have $\tilde{f}_1(a) = \tilde{f}_2(a)$ since the homeomorphism $\pi : V \rightarrow U$ maps them both to $f(a)$. This shows that the set $\{w \in W \mid \tilde{f}_1(w) = \tilde{f}_2(w)\}$ is open. But this set is also closed since it is the inverse image of the diagonal under the map $\tilde{f}_1 \times \tilde{f}_2 : W \rightarrow E \times E$, and the diagonal is closed since E is Hausdorff. Since W is connected, this set is either empty or all of W . ■

Corollary 4.2.3 ([19, Cor. 4.5]). Let $\pi_i : E_i \rightarrow M$, $i = 1, 2$, be covering maps such that E_1 is simply connected, and let $p_i \in E_i$ be such that $\pi_1(p_1) = \pi_2(p_2)$. Then there is a unique morphism of covering spaces $f : E_1 \rightarrow E_2$ such that $f(p_1) = p_2$. Moreover, f is a covering map.

Proof. By the Lifting Theorem 4.1.18, since $\pi_1(E_1, p_1) = 1$, there is a lifting of $\pi_1 \rightarrow E_1 \rightarrow M$ to a map $f : E^1 \rightarrow E^2$ such that $\pi_2 \circ f = \pi_1$. It is unique by Lemma 4.2.2. It is an easy exercise to show that f is a covering map. ■

Corollary 4.2.4 ([19, Cor. 4.6]). *If in Corollary 4.2.3 both E_1 and E_2 are simply connected, then the unique map f is a homeomorphism.*

Proof. By Corollary 4.2.3, in addition to f there is also an analogous map $k : E_2 \rightarrow E_1$. Then $k \circ f : E_1 \rightarrow E_1$ covers the identity map at p_1 . By Lemma 4.2.2, it must equal the identity everywhere. Similarly $g \circ k = \text{id}_{E_2}$, hence $k = f^{-1}$. ■

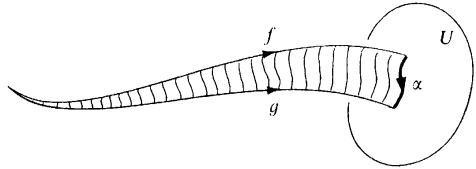
This implies that all simply connected covering spaces of a given base space M are isomorphic in the category $\text{Cov}_{\bullet M}$. Such covering spaces are called universal, and are guaranteed to exist under a mild restriction.

Definition 4.2.5 (Locally relatively simply connected space). A space X is locally relatively simply connected (or semilocally simply connected) if each point has a neighborhood U such that all loops inside of U are homotopically trivial in X (i.e. for any $u \in U$ the homomorphism $\pi_1(U, u) \rightarrow \pi_1(X, u)$ is trivial).

One can show that the existence of a simply connected covering space is equivalent to M being locally relatively simply connected (see [19, Theorem 8.4]). Since we are mostly interested in manifolds, we will only prove the existence. Let us only work with connected, locally simply connected spaces (in particular, connected manifolds).

Theorem 4.2.6 (Universal cover). *For a given base M that is a connected and locally relatively simply connected Hausdorff space, there is a unique (up to homeomorphism) topological covering space $E \xrightarrow{\pi} M$ whose total space E is simply connected. This covering space is denoted by \tilde{M} and is called the universal cover of M .*

Proof. See [8, Thm 11.43] or [19, Thm 8.4]. The basic idea is to explicitly construct \tilde{M} as the space of endpoints of all possible paths in M based at a fixed point $m \in M$. Two paths are equivalent if their endpoints coincide and they are homotopic with fixed endpoints. The topology on this set of paths is introduced by considering sets of homotopic paths with endpoints belonging to an open set in M (and the homotopies fix the starting point but can move the endpoint within the open set, see figure). The projection π maps a path to its endpoint. It is not difficult to check that this is indeed a topological covering map and the total space is simply connected. Uniqueness will follow from the universal property described below.



Remark 4.2.7. By examining the details of the proof one can notice that local simply-connectedness is not actually required, and semilocal simply-connectedness is sufficient. The converse direction is clear: if a loop is in an evenly covered subspace of M then it lifts to E , and if E is simply connected, the loop must be trivial in E . But then the composition of its homotopy to a constant loop with the covering map π shows that the original loop in M must be trivial. Thus we have shown that a connected and locally path-connected M has a universal covering space iff it is locally relatively simply connected.

Example 4.2.8. The Hawaiian earring (Example 2.5.30), as well as the infinite-dimensional torus $\mathbb{T}^\infty = \prod_{i=1}^\infty \mathbb{S}^1$, are not locally relatively simply connected, and thus don't have universal coverings.

Theorem 4.2.9. Let $E \xrightarrow{\pi} M$ be a topological covering map and fix a smooth structure on M . Then E is a topological manifold that supports a unique smooth structure such that π is a smooth covering map with $\pi^{-1}(\partial M) = \partial E$.

Proof. See [7, Prop 4.40].

This theorem essentially means that in order to classify smooth covering spaces above a smooth base manifold, it suffices to only classify the topological covering spaces, and they are in one-to-one correspondence with the smooth ones. In other words, the theory of covering spaces is an intrinsically topological theory that doesn't gain or lose anything from imposing smooth structures.

Corollary 4.2.10. For a connected smooth manifold M , there exists a unique simply connected smooth covering space called the universal covering manifold of M .

From now on we won't really distinguish between smooth and topological covering spaces, although we only care about the smooth case.

Corollaries 4.2.3 and 4.2.4 can now be restated in terms of morphisms:

Corollary 4.2.11 (Universality of the universal cover).

1. Any morphism of covering spaces is itself a covering map;
2. If M is a topological manifold (or any space that has a universal covering space), then \tilde{M} is the initial object in the category $\text{Cov}_{\bullet M}$ of pointed covering spaces of M (i.e. objects are $E \xrightarrow{\pi} M$ with a basepoint $p_0 \in E$ and morphisms must respect the basepoints).

In the category of non-pointed covering spaces, the universal cover is not an initial object, but the arrows coming out of it are unique up to automorphisms of the two covering spaces at the ends of the arrow. We now describe the sets of such automorphisms.

4.3 Deck transformations

Definition 4.3.1 (Group action). An (right/left) action of a group G on a set (or a space, a manifold) X is a map $\Phi : G \times X \rightarrow X$, with its partial maps denoted by $\Phi_g = \Phi(g, _) : X \rightarrow X$ and $\Phi_x = \Phi(_, x) : G \rightarrow X$, such that $\Phi_e = \text{id}_X$ and one of the following:

$$\text{left action : } \Phi_g \circ \Phi_h = \Phi_{gh}, \quad (4.3.1)$$

$$\text{right action : } \Phi_g \circ \Phi_h = \Phi_{hg}. \quad (4.3.2)$$

In other words, right actions are the same thing as left actions on the opposite group G^{op} . For this reason right actions are usually written as $x \cdot g = \Phi_g(x)$ and left actions as $g \cdot x = \Phi_g(x)$. The sets $G(x) = \Phi_x(G)$ are called *orbits*.

In the topological and smooth categories one usually considers continuous or smooth group actions, provided G is a Lie group, but we will delay that discussion until Section 6.6. For now G will always be a discrete group (i.e. a group with discrete topology), so any condition of continuity or smoothness on the action is simply a condition on each Φ_g individually.

Definition 4.3.2 (G -equivariant maps). The category of sets (spaces, manifolds) carrying a (left) action of G is called G -Set. Morphisms in this category are maps $f : X \rightarrow Y$ such that for all $g \in G$, $f(g \cdot x) = g \cdot f(x)$. These maps are called G -equivariant. Similarly, G -Set $_{\bullet}$ consists of pointed sets with equivariant maps that respect basepoints.

Definition 4.3.3 (Stabilizer). Given a group G acting on a set X , and a point $x \in X$, the stabilizer of x is defined as the subgroup $G_x = \{g \in G \mid \Phi(g, x) = x\}$.

Definition 4.3.4 (Transitive action). An action Φ is called transitive if it has only one orbit, i.e. for each pair $x, y \in X$ there is an $g \in G$ such that $\Phi(g, x) = y$.

Definition 4.3.5 (Simply transitive action). An action Φ is called simply transitive if for each pair $x, y \in X$ there is a *unique* element $g \in G$ such that $\Phi(g, x) = y$.

Definition 4.3.6 (Faithful action). An action Φ is called faithful if no nontrivial elements of G act trivially: $\Phi_g = \text{id}_X \Leftrightarrow g = e$.

Definition 4.3.7 (Free action). An action Φ is called free if no nontrivial elements of G have any fixed points: $\exists x \in X : \Phi(g, x) = x \Leftrightarrow g = e$.

Proposition 4.3.8. *An action is simply transitive iff it is free and transitive.*

Proof. Assume a left action for simplicity. For the necessary condition, we only need to show freeness. If $\Phi_g = \text{id}_X$, then $g \cdot x = x = e \cdot x$, therefore by simple transitivity $g = e$.

For the converse, assume there are two elements such that $g_1 \cdot x = g_2 \cdot x = y$. Then $g_2^{-1}g_1 \cdot x = x$, which means $g_2^{-1}g_1 = e$ by freeness. ■

Throughout this section let $E \xrightarrow{\pi} M$ be a covering space with a fixed point $x_0 \in M$. For simplicity we also denote

$$G = \pi_1(M, x_0), \quad F = \pi^{-1}(x_0). \quad (4.3.3)$$

The discrete set F is called the *fiber* of π over x_0 . We are going to describe an action of the group G on F as a group of permutations.

Let $p \in F$ and let $g \in G$ be an element represented by a loop $\gamma : I \rightarrow M$. Lift it to get a map $\tilde{\gamma}$ in E with $\tilde{\gamma}(0) = p$. Then define

$$\Phi(g, p) = p \cdot g = \tilde{\gamma}(1) \in F. \quad (4.3.4)$$

By Corollary 4.1.10 this does not depend on the choice of γ and thus is a well-defined map $\Phi : F \times G \rightarrow F$. Now we can check some properties of this function:

1. $x \cdot e = x$.
2. $(x \cdot g) \cdot h = x \cdot (gh)$, i.e. it is a right action (because the product of two loops gh corresponds to the application of Φ_g and then Φ_h).
3. E is connected iff this action is transitive (i.e. there is only one orbit, $p \cdot G = F$). To see this, pick a path in E from p_0 to another $p \in F$. This projects onto a loop γ in M and $g = [\gamma]$ provides the required element: $p_0 \cdot p = p$. Conversely, if E is not connected, the action is obviously not transitive.
4. The stabilizer of a point is $G_{p_0} = \{g \in G \mid p_0 \cdot g = p_0\}$. Then

$$G_{p_0} = \text{im } \pi_* = \pi_*(\pi_1(E, p_0)) \quad (4.3.5)$$

where $\pi_* : \pi_1(E, p_0) \rightarrow \pi_1(M, x_0) = G$ is the homomorphism induced by π . This is true because $g \in G_{x_0}$ iff $g = [\gamma]$ and γ lifts to a loop at p_0 , i.e. iff $g \in \text{im } p_*$.

5. Assuming E is connected, the map $\phi : G_{p_0} \setminus^G \rightarrow F$ taking the right coset $G_{x_0} g$ to $p_0 \cdot (G_{p_0} g) = p_0 \cdot g$ is a bijection.

Obviously any covering space decomposes into a disjoint union of connected components, and the action of $\pi_1(M)$ is transitive on the fiber of each component separately. In what follows, we generally assume that E is connected. In summary, we have the following Theorem.

Theorem 4.3.9. *Given a connected covering space $\pi : E \rightarrow M$, there is a one-to-one correspondence between the set $\pi_*(\pi_1(E, p_0)) \setminus^{\pi_1(M, x_0)}$ of right cosets, and the fiber $\pi^{-1}(x_0)$. (Note that $\pi_*(\pi_1(E, p_0)) \cong \pi_1(E, p_0)$ since π_* is a monomorphism by Corollary 4.1.12.)*

Corollary 4.3.10. *The number of sheets of a connected covering space equals the index (number of left or right cosets) of $G_{p_0} = \pi_*(\pi_1(E, p_0))$ in $G = \pi_1(M, x_0)$.*

Corollary 4.3.11. *If $\tilde{\pi} : \tilde{M} \rightarrow M$ is a universal covering map (i.e. \tilde{M} is simply connected), then the number of sheets equals the order (cardinality) of the group $\pi_1(M, p_0)$.*

Example 4.3.12 (Fundamental groups of real projective spaces). Since \mathbb{S}^n is simply connected for $n > 1$ and \mathbb{S}^n is a double covering of the real projective space $\mathbb{R}P^n$, it follows that $\pi_1(\mathbb{R}P^n) \cong \mathbb{Z}_2$.

Definition 4.3.13. For a covering space $E \xrightarrow{\pi} M$, define the group $\text{Aut}(\pi)$ of automorphisms of this covering space in the category Cov_M (also called *deck transformations*). Its elements are homeomorphisms of E that cover the identity map on M . For a given elementary open set $U \subset M$, they permute the connected components U_i of the pre-image $\pi^{-1}(U)$. In other words, the group $\text{Aut}(\pi)$ acts on E via morphisms of covering spaces.

Proposition 4.3.14 ([19, Prop. 6.2]). *If $a \in \text{Aut}(\pi)$ and $g \in \pi_1(M, p_0)$ and $x \in \pi^{-1}(x_0)$ then $a(x) \cdot g = a(x \cdot g)$. In other words, deck transformations are equivariant with respect to the action of $\pi_1(M, x_0)$.*

Proof. Let γ be a loop at x_0 representing g and lift it to a path $\tilde{\gamma}$ starting at p . Then $\tilde{\gamma}(1) = p \cdot g$ by definition. Look at the path $a \circ \tilde{\gamma}$. It is a lifting of γ and starts at $a(p)$. Thus it ends at $a(p) \cdot g$ by definition of the latter. But it also ends at $a(\tilde{\gamma}(1))$, which is $a(p \cdot g)$. ■

This proof actually shows that not only automorphisms are G -equivariant, but all morphisms in the category Cov_M are. This means that the action of $\pi_1(M)$ that we've constructed is *natural*. These results can be summarized by the following Theorem.

Theorem 4.3.15 (Monodromy Action). *For any covering space $E \xrightarrow{\pi} M$,*

1. *there is a natural homomorphism $\pi_1(M) \rightarrow \text{Aut}(\pi)$ called the monodromy action;*
2. *$\text{Aut}(\tilde{\pi}) \cong \pi_1(M)$, where $\tilde{\pi}$ is the universal cover.*

Definition 4.3.16 (Normalizer). For a subset (not necessarily a subgroup) $H \subset G$ of a group G , its normalizer is

$$N_G(H) = \{n \in G \mid nH = Hn\}. \quad (4.3.6)$$

A subgroup H is normal in G iff $N_G(H) = G$. Otherwise, $N_G(H)$ is the largest subgroup of G such that H is a normal subgroup of $N_G(H)$.

Remark 4.3.17. If H is a subgroup, then $H \subset N_H(H)$. Moreover, $N_G(g^{-1}Hg) = g^{-1}N_G(H)g$, so the normalizers of conjugate subsets are conjugate.

The following lemma is the analog of the “change of coordinates” formulas in linear algebra: mapping the coordinates of vectors to a new basis is equivalent to conjugating the matrices representing linear operators.

Lemma 4.3.18. *For any group G acting on a set X from the right,*

$$G_{p_0 \cdot g} = g^{-1}G_{p_0}g. \quad (4.3.7)$$

Similarly, for left actions, $G_{g \cdot p_0} = gG_{p_0}g^{-1}$.

Proof.

$$G_{p_0 \cdot g} = \{h \mid (p_0 \cdot g) \cdot h = p_0 \cdot g\} = \{h \mid p_0 \cdot ghg^{-1} = p_0\} = \{h \mid ghg^{-1} \in G_{p_0}\}, \quad (4.3.8)$$

■

The most important question in the study of deck transformations is: given two points in the fiber, when does there exist a deck transformation mapping one to the other? This is answered in the following theorem.

Theorem 4.3.19 ([19, Thm. 6.3]). *Provided E is connected, let $p, p_0 \in F = \pi^{-1}(x_0)$, $G = \pi_1(M, x_0)$ and $G_q = \pi_*(\pi_1(E, q))$ for any $q \in E$. Then the following are equivalent (TFAE):*

1. $\exists a \in \text{Aut}(\pi)$ such that $a(p_0) = p$.
2. $\exists g \in N_G(G_{p_0})$ such that $p_0 \cdot g = p$.
3. $G_{p_0} = G_p$.

Proof. Recall that $G_{p_0} = \pi_*(\pi_1(E, p_0))$ and $G_p = \pi_*(\pi_1(E, p))$. By the Lifting Theorem 4.1.18 a map a covering the identity and taking p_0 to p exists iff $\pi_*(\pi_1(E, p_0)) \subset \pi_*(\pi_1(E, p))$. Similarly, a map a' covering the identity and mapping p to p_0 exists iff the opposite inclusion holds. If both exist then $a \circ a'$ covers the identity and has a point in common with the identity, therefore $a \circ a' = \text{id}_E$ by Lemma 4.2.2. This proves the equivalence (1) \Leftrightarrow (3).

Now we can prove (2) \Rightarrow (3): if $p = p_0 \cdot g$ and $g \in N(G_{p_0})$ then $G_p = G_{p_0 \cdot g} = g^{-1}G_{p_0}g = G_{p_0}$ as claimed.

For (3) \Rightarrow (2), suppose $G_{p_0} = G_p$ and $p = p_0 \cdot g$ (such a g exists since G acts transitively on F). Then $G_{p_0} = G_p = G_{p_0 \cdot g} = g^{-1}G_{p_0}g$, so $g \in N(G_{p_0})$. ■

From (2) \Leftrightarrow (1) and the last part of the proof we have the following Corollary.

Corollary 4.3.20 ([19, Cor. 6.4]). *For a connected covering space, the subgroup $G_{p_0} = \pi_*(\pi_1(E, p_0))$ is normal in $G = \pi_1(M, x_0)$ iff the action of $\text{Aut}(\pi)$ is (simply) transitive on $F = \pi^{-1}(x_0)$.*

Corollary 4.3.21 ([19, Cor. 6.5]). *In a connected covering space, if p ranges over the fiber F then G_p ranges over all conjugates of $G_{p_0} = \pi_*(\pi_1(E, p_0))$.*

Proof. This is a consequence of (4.3.7). ■

Remark 4.3.22. It is worth providing some intuition for normalizers. If two points in a G -set belong to the same orbit, i.e. they can be mapped into each other by some elements of G , then their stabilizers are obviously conjugate subgroups of G .

By definition, conjugation doesn't change a normal subgroup (as a set, not pointwise). Therefore if a stabilizer of a point is normal, then all other points in that orbit have the same exact stabilizer. In particular, this means that the only elements of the group which actually move the points in the orbit are the cosets of the normal stabilizer. Moreover, if you have two group elements from the same coset of this normal stabilizer, then they act on the orbit *in the exact same manner*.

Definition 4.3.23 (Regular covering map). A covering map π is called regular if $\text{Aut}(\pi)$ is transitive on the fiber. In the case of a connected total space this is equivalent to $G_{p_0} = \pi_*(\pi_1(E, p_0))$ being normal in $\pi_1(M, x_0)$.

In other words, normal subgroups of $\pi_1(M, x_0)$ correspond to regular connected covering spaces of M . A regular covering space is one whose fundamental group “looks the same” in terms of the generators of $\pi_1(M, x_0)$ regardless of the choice of the basepoint p_0 in the fiber. Such covering spaces have an inherent “symmetry” given by the factor G/G_{p_0} (which exists

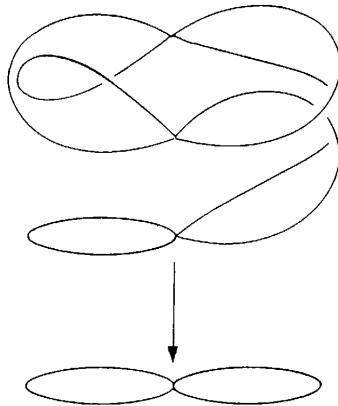


Fig. 4.2: This threefold cover of the figure eight space is not regular. Indeed, its group of automorphisms is trivial.

because G_{p_0} is normal). Obviously any covering space that is not connected can't be regular. For an example of a connected non-regular covering space see Figure 4.2.

Remark 4.3.24. These ideas comprise the powerful *Reidemeister-Schreier method* for finding presentations of subgroups $H \leq G$ given a presentation of G . Namely, one constructs a base space whose fundamental group is G , then finds a covering space corresponding to H and computes its fundamental group.

Definition 4.3.25. Define the map $\Theta : N_G(G_{p_0}) \rightarrow \text{Aut}(\pi)$ by $\Theta(g) = a_g$ where a_g is the unique deck transformation such that $a_g(p_0) = p_0 \cdot g$.

Theorem 4.3.26 (Classification of Deck Transformations [19, Thm. 6.8]). *$\ker \Theta = G_{p_0}$, and consequently*

$$\text{Aut}(\pi) \cong N_G(G_{p_0})/G_{p_0}. \quad (4.3.9)$$

Proof. First,

$$a_h(a_g(p_0)) = a_h(p_0 \cdot g) = a_h(p_0) \cdot g = (p_0 \cdot h) \cdot g = p_0 \cdot (hg) = a_{hg}(p_0). \quad (4.3.10)$$

This proves that Θ is a homomorphism. Next note that if $a \in \text{Aut}(\pi)$ then there is an $g \in N(G_{p_0})$ such that $a(p_0) = p_0 \cdot g = a_g(p_0)$. Therefore $a = a_g$, which shows the surjectivity of Θ .

Finally we compute the kernel of Θ . The result follows from the chain of equivalences

$$a_g = \text{id}_E \Leftrightarrow p_0 \cdot g = p_0 \Leftrightarrow g \in G_{p_0}. \quad (4.3.11)$$

■

Corollary 4.3.27. *If the covering map $E \xrightarrow{\pi} M$ is regular, then*

$$\text{Aut}(\pi) \cong G/G_{p_0} \quad (4.3.12)$$

Corollary 4.3.28. *If $\tilde{M} \xrightarrow{\tilde{\pi}} M$ is a universal covering map then*

$$\text{Aut}(\tilde{\pi}) \cong G = \pi_1(M, x_0). \quad (4.3.13)$$

4.4 Classification of covering spaces

Our examination of connected covering spaces allows us to easily prove the following classification of connected covering spaces in terms of subgroups of the fundamental group of the base.

Theorem 4.4.1 (Classification of covering maps I). *Let M be path-connected and semilocally simply connected so that it has a universal covering space \tilde{M} . Then the following Galois correspondence holds:*

1. *The isomorphism classes of connected pointed covering spaces of M are in one-to-one correspondence with subgroups of $G = \pi_1(M, x_0)$.*
 2. *The isomorphism classes of connected covering spaces of M without basepoints are in one-to-one correspondence with the conjugacy classes of subgroups of $G = \pi_1(M, x_0)$.*
- The correspondence is given by $E \leftrightarrow \pi_*(\pi_1(E, p_0)) = G_{p_0}$ where π is the covering map.*

Proof. The second statement follows from the first combined with Corollary 4.3.21. It remains to show that the map taking a covering map with a basepoint, $\pi : (E, p_0) \rightarrow (M, x_0)$, into the subgroup $\pi_*(\pi_1(E, p_0))$ of $\pi_1(M, x_0)$ is a bijection. It is an injection by the Lifting Theorem 4.1.18.

To see that it is a surjection, suppose that $H \subset \pi_1(M, x_0) = G$ is a subgroup. Since \tilde{M} is simply connected, Theorem 4.3.26 gives the isomorphism $\Theta : G \rightarrow \text{Aut}(\pi)$, where $\Theta(g) = a_g$. Under this map, H is mapped to a subgroup $A_H \leq \text{Aut}(\pi)$. Put $E = \tilde{M}/A_H$ which projects to M canonically. Let p_0 be the image of the basepoint $\tilde{p}_0 \in \tilde{M}$. We wish to identify $\pi_*(\pi_1(E, p_0))$. Let γ be a loop in X at p_0 . Lifting this to \tilde{M} at \tilde{p}_0 gives the same path as a lifting of the projection of γ to a loop at x_0 in M . Thus the lifting ends at $a_g(\tilde{p}_0)$,

where $g \in \pi_1(M, x_0)$ is the homotopy class of the projection of γ to M . But for γ to be a loop in $E = \tilde{M}/A_H$, we must have that \tilde{p}_0 and $a_g(\tilde{p}_0)$ are in the same orbit of A_H . This is true iff $a_g \in A_H$, and this holds iff $g \in H$. But g is an arbitrary element of $\pi_*(\pi_1(E, p_0))$. Thus $H = \pi_*(\pi_1(E, p_0))$. ■

Now we restate the classification theorem as it was described in Example 1.7.16. We consider all, not only connected, coverings.

Given an action of a group G on a space X , the space X decomposes into a disjoint union of orbits of G . Let us assume a discrete topology on G , in which case a continuous/smooth action is simply one for which the maps $\Phi_g : X \rightarrow X$ are continuous/smooth for all g . The set of all orbits is denoted by X/G and, if imbued with the quotient topology descended from X , is called the *orbit space*. Note that the canonical projection $p : X \rightarrow X/G$ that takes a point to its orbit is open because for an open $U \subset X$, $p^{-1}(p(U)) = \bigcup_g (g \cdot U)$, which is a union of open sets.

Every covering map of M gives rise to an action of $G = \pi_1(M, x_0)$ on the fiber $F = F_{x_0}$. Under this action F decomposes into a disjoint union of orbits F_j , each of them associated with a conjugacy class of stabilizer subgroups $H_j < G$. This determines a functor from the category Cov_M of coverings of M to the category $G\text{-Set}$ of sets with a (left) action of $G = \pi_1(M, x_0)$. If F is a G -set with a transitive action of G , then the stabilizers of any two of its points are conjugate subgroups of G . Choosing $p_0 \in F$ and denoting $H = G_{p_0}$, we have that F is isomorphic to the set of cosets G/H , also called a *homogeneous set*. From such a G -set F we can easily reconstruct a connected covering space π by taking the universal covering space \tilde{M} and letting $\pi : \tilde{M} \rightarrow \tilde{M}/H$ be the orbit map of the action of H on \tilde{M} . If F is not homogeneous, then it is a disjoint union of homogeneous sets (orbits), and the corresponding covering space is a disjoint union of the corresponding connected ones. This provides a functor in the opposite direction. It is easy to check that these two functors are each other's quasi-inverses. We thus have the second version of the theorem.

Theorem 4.4.2 (Classification of coverings II [26, Thm. 3.3.2]). *If M admits a universal covering space, then the category Cov_M is equivalent to the category $\pi_1(M)\text{-Set}$. The corresponding functor returns the action of $\pi_1(M)$ on the fiber of the covering space and restricts morphisms to the fiber. Its quasi-inverse reconstructs a connected component of a covering space from each orbit of the $\pi_1(M)$ -set as the quotient of the universal covering by the action of the stabilizer subgroup. Similarly, $\text{Cov}_{M\bullet}$ is equivalent to $\pi_1(M)\text{-Set}$.*

Remark 4.4.3. An action of a group G on a set X is really a homomorphism from G into the group $\text{Aut}(X)$ of automorphisms (bijective mappings, a.k.a. permutations) of X . The element of $\text{Hom}(\pi_1(M), \text{Aut}(F))$ representing a given covering space is called the *characteristic class* of this covering space. Therefore for a given discrete fiber F , all covering spaces of M

with fiber F are completely classified by a characteristic class that is a contravariant functor from the category $\text{Cov}_M(F)$ of all such covering spaces to the category of $\pi_1(M)$ actions on F .

The next Corollary follows from the fact that the fundamental group π_1 is a homotopy invariant.

Corollary 4.4.4. *The isomorphism classes of covering spaces of two base manifolds that are homotopy equivalent are in one-to-one correspondence. That is, the set of (isoclasses of) covering spaces is a homotopy invariant.*

Remark 4.4.5. Due to the path lifting property, we have a set-valued *transport functor* $T_\pi : \Pi(M) \rightarrow \text{Set}$ on the fundamental groupoid that takes a path γ in M , say connecting $x \in M$ to $y \in M$, and produces a map $T_\pi[\gamma] : F_x \rightarrow F_y$. Thus T_π is an element of the category $[\Pi(M), \text{Set}]$ consisting of functors from $\Pi(M)$ to Set and natural transformations between them. A morphism of covering spaces also induces a natural transformation between their transport functors (covariantly), therefore we have a functor $T : \text{Cov}_M \rightarrow [\Pi(M), \text{Set}]$. The classification theorem can be restated in yet another form: if M is path-connected and semilocally simply connected, then T is an equivalence of categories.

Now we focus on regular coverings and to each such covering associate a functor that produces other (non-regular) coverings, providing another valuable version of the classification theorem.

Definition 4.4.6 (Properly discontinuous action). An action of a group G on a space X is properly discontinuous if each point $x \in X$ has a neighborhood U such that $\exists g \cdot U \cap U \neq \emptyset \Leftrightarrow g = e$.

Remark 4.4.7. Any properly discontinuous action is free. Indeed, suppose $g \cdot x = x$ for some $g \in G$ and $x \in X$. Then the neighborhood U of x guaranteed to exist by definition of a properly discontinuous action necessarily satisfies $g \cdot U \cap U \neq \emptyset$, and therefore $g = e$. Hence the action is free.

Definition 4.4.8 (G -principal covering). A right G -principal covering is a covering map $\pi : E \rightarrow B$ with a properly discontinuous action of G on E such that $\pi(p \cdot g) = \pi(p)$ for all $(g, p) \in G \times E$ and such that the induced action on each fiber is transitive.

Example 4.4.9. The orbit map $E \rightarrow E/G$ of any properly discontinuous action is obviously a G -principal bundle. Conversely, every G -principal covering $\pi : E \rightarrow B$ induces a homeomorphism $\pi' : E/G \cong B$ given by $p \cdot G \mapsto \pi(p)$. Indeed, since the action is transitive, the fiber $\pi^{-1}(x)$ coincides with the orbit $p \cdot G$ for any $p \in \pi^{-1}(x)$. Hence an inverse to π' is given by $x \mapsto \pi^{-1}(x) = p \cdot G$. Its continuity follows from the fact that π is a local homeomorphism.

We now examine the relationship between principal and regular coverings.

Proposition 4.4.10 ([26, Prop. 3.1.8]). *Let $\pi : E \rightarrow B$ be a covering map. Then:*

- (a) *If E is connected, then the action of $\text{Aut}(\pi)$ (or any of its subgroups) is properly discontinuous.*
- (b) *Let B be locally path connected and let H be a subgroup of $\text{Aut}(\pi)$. Then the map $\pi' : E/H \rightarrow B$ induced by π is a covering map.*

Proof. (1) Let $p \in E$ and $g \in \text{Aut}(\pi)$. Let U be an evenly covered neighborhood of $\pi(p)$, and let U_p be the sheet over U containing p . For $q \in U_p \cap gU_p$ we have $\pi(q) = \pi(g^{-1}q)$, since g^{-1} is an automorphism. Hence $q = g^{-1}q$, since both elements are contained in U_p . This shows $g^{-1} = e$, hence $U_p \cap gU_p = \emptyset$ for $g \neq e$, and we see that the action is properly discontinuous.

(2) Let $U \subset B$ be open, path connected, and evenly covered. Let $\pi^{-1}(U) = \bigcup_j U_j$ be the decomposition into sheets over U . An element $h \in H$ permutes the sheets, since they are the path components of $\pi^{-1}(U)$. The equivalence classes with respect to H are therefore open in the quotient topology of E/H and are mapped bijectively and continuously under π' . Since π is open, so is π' . Hence π' is trivial over U . Since B is locally path connected, it has an open covering by such sets U . ■

Corollary 4.4.11. *A G -principal covering $\pi : E \rightarrow B$ induces a homeomorphism $E/G \cong B$.*

Thus the group $\text{Aut}(\pi)$ of deck transformations of any covering space $E \xrightarrow{\pi} M$ acts properly discontinuously, and if π is regular, this action is transitive on each fiber. Therefore

a regular covering π is an $\text{Aut}(\pi)$ -principal covering.

Now we ask the opposite question: what kind of coverings can properly discontinuous actions give rise to?

Proposition 4.4.12. *If G acts properly discontinuously on a path-connected and locally path-connected Hausdorff space E , then the orbit map $\pi : E \rightarrow E/G$ is a regular covering map with deck transformation group $\text{Aut}(\pi) = G$.*

Proof. Let $U \subset X$ be a path-connected open set as in the definition of properly discontinuous actions and put $U^* = \pi(U)$, which is open as remarked above. Since $U \rightarrow U^*$ is continuous, U^* is path-connected. Also, the sets $g \cdot U$ are the components of $\pi^{-1}(U^*)$. The maps $g \cdot U \rightarrow U^*$ are continuous, open, injective and surjective, and hence homeomorphisms. Thus π is a covering map. Elements of G are deck transformations and act transitively on the fiber. There are no other deck transformations by Lemma 4.2.2. ■

Corollary 4.4.13. *If X is simply connected and locally path-connected and G acts properly discontinuously on X , then $\pi_1(X/G) \cong G$.*

Now let π be a G -principal covering, so that π induces a homeomorphism $E/G \cong B$, and by the above Proposition $\text{Aut}(\pi) \cong G$. Therefore every principal covering is regular. We conclude that

a G -principal covering π is a regular covering with group of deck transformations G .

We now explain how, given a regular (or principal) covering, one can produce other coverings.

Definition 4.4.14 (Associated covering). A right G -principal covering $\pi : E \rightarrow B$ gives rise to associated coverings. Let F be a set with left G -action (a G -set). Denote by $E \times_G F$ the quotient space $E \times F / \sim$ under the equivalence relation $(p, f) \sim (p \cdot g^{-1}, g \cdot f)$ for $(p, f, g) \in E \times F \times G$. The continuous map $\pi_F : E \times_G F \rightarrow B$, $(p, f) \mapsto \pi(p)$ is a covering with fiber F .

A G -equivariant map $\psi : F_1 \rightarrow F_2$ induces a morphism of coverings

$$\text{id}_{E \times_G F} \circ \psi : E \times_G F_1 \rightarrow E \times_G F_2, (p, f) \mapsto (p, \psi(f)). \quad (4.4.1)$$

Thus for every right G -principal covering π (over some base B) we have constructed the “associated covering functor” $A(\pi) : G\text{-Set} \rightarrow \text{Cov}_B$ from the category of left G -sets (and equivariant maps) to the category of covering spaces of B .

Definition 4.4.15 (Universal G -principal covering). We call a G -principal covering π *universal* (among all G -principal coverings over all possible base spaces) if the functor $A(\pi)$ is an equivalence of categories.

The classification theorem 4.4.2 can now be restated as follows.

Theorem 4.4.16 (Classification of covering maps III). *If a covering map $\pi : E \rightarrow B$ is universal (i.e. E is simply connected), then it is also universal $\pi_1(B)$ -principal (i.e. the functor $A(\pi)$ is an equivalence of categories).*

In this sense, universal coverings are “classifying objects” for all other coverings. Now it is natural to ask whether the converse holds, i.e. whether every universal G -principal covering is the universal covering of some space B such that $\pi_1(B) \cong G$. For this it suffices to show that any universal G -principal covering must be simply connected.

Proposition 4.4.17. *Let B be a space which admits a universal covering. If a right G -principal covering $\pi : E \rightarrow B$ is universal, then E is simply connected, i.e. π is a universal covering of B .*

Proof. First consider the group G acting on itself from the left as a G -set. It is clear from the definition that the associated bundle $A(\pi)(G)$ is isomorphic to π itself. Now consider any other covering $\pi' : E' \rightarrow B$. By assumption, $A(\pi) : G\text{-Set} \rightarrow \text{Cov}_B$ is an equivalence of categories, thus there exists a left G -set F such that the covering $A(\pi)(F)$ is isomorphic to π' . Now pick a point $p_0 \in F$. The action $G \curvearrowright F$ induces a G -equivariant map $g \mapsto g \cdot p_0$. This is a morphism $f : G \rightarrow F$ in the category of G -sets. Since $A(\pi)$ is a functor, $A(\pi)(f)$ is a morphism from $A(\pi)(G)$ to $A(\pi)(F)$. Since $A(\pi)$ is an equivalence of categories, there is a natural transformation that turns this morphism into a morphism from π to π' . Therefore π is an initial object in the category Cov_B , and hence isomorphic to the universal covering. ■

We thus have the following abstract reformulation of the classification theorem, which will be a useful reference point for when we classify general fiber bundles.

Theorem 4.4.18 (Classification of covering maps IV [26, Thm. 3.4.3]). *A covering map $\pi : E \rightarrow B$ is universal G -principal for some group G iff it is a universal covering map of B and $G \cong \pi_1(B)$.*

Note that for a fixed G there are many non-isomorphic universal G -principal coverings: the universal covering of any space B whose fundamental group is G will work.

Remark 4.4.19. We will later learn to construct, for any discrete group G , a space X_G whose fundamental group is isomorphic to G and which admits a universal covering. Therefore for every group G there exists a universal G -principal covering, although it is not unique. Once we learn about CW-complexes we will be able to define a special universal G -principal covering that is unique up to homotopy (namely the universal covering of the first Eilenberg-MacLane space $K(G, 1)$, see Section 5.7). This fact will eventually have a natural extension to all groups and fiber bundles: for any group G we will construct a so called classifying space BG and a principal G -bundle $EG \rightarrow BG$ such that EG is contractible. There will hold isomorphisms $\pi_{n+1}(BG) \cong \pi_n(G)$ for all $n \geq 0$. In the case of discrete G the bundle $EG \rightarrow BG$ is obviously a covering space, and the aforementioned space X_G provides a universal covering that is an “approximation” of this space, and for which only the first isomorphism $\pi_1(B) \cong \pi_0(G) \cong G$ holds.

Since every covering space is a disjoint union of its connected components, let us separately address the classification of *connected* covering spaces in terms of equivalence of certain categories. Let $\pi : E \rightarrow B$ be a universal right G -principal covering. A left G -set A is a disjoint sum of its orbits. We thus have a corresponding decomposition of the total space $E \times_G F$ into the sum of $E \times_G C$, where C runs over the orbits of F . An orbit is a transitive G -set and is isomorphic to a homogeneous set (set of cosets) G/H of some subgroup $H < G$. The homeomorphism $E \times_G G/H \cong E/H$ shows that the summands

$E \times_G C$ are path-connected. The action of H on E is properly discontinuous and therefore $E \rightarrow E/H$ is an H -principal covering. Also the induced map $\pi_H : E/H \rightarrow B$ is a covering.

The category of homogeneous G -sets and G -equivariant maps is called the *orbit category* $\text{Orb}(G)$. The sets G/K and G/L are isomorphic iff the subgroups K and L are conjugate in G . The stabilizers of points of G/H are conjugate to H . The inclusion of the subcategory $\text{Orb}(G)$ into the category of transitive G -sets is an equivalence.

Let $\tilde{\pi} : \tilde{E} \rightarrow B$ be a universal right G -principal covering with simply connected E . Then the functor $A(\tilde{\pi})$ induces an equivalence of $\text{Orb}(G)$ with the category of connected coverings of B . Each connected covering is thus isomorphic to a covering of the form $\pi_H : \tilde{E}/H \rightarrow B$ for a subgroup $H < G$.

Let $\pi : E \rightarrow B$ be a connected covering. The subgroup $\pi_*(\pi_1(E, p_0)) \subset \pi_1(B, x_0)$ is called the characteristic subgroup $C(\pi, p_0)$ of π with respect to p_0 . Different basepoints p_0 lead to conjugate characteristic subgroups, and conversely, every subgroup conjugate to $C(\pi, p_0)$ arises this way. We summarize these results in the final version of the classification theorem:

Theorem 4.4.20 (Classification V [26, Thm. 3.4.4]). *If M admits a universal covering space, then the category of connected coverings of M is equivalent to the orbit category $\text{Orb}(\pi_1(M, x_0))$. The isomorphism class of a connected covering $\pi : E \rightarrow M$ corresponds under this equivalence to the isomorphism class of $\pi_1(M, x_0)/C(\pi, p_0)$ for any $p \in \pi^{-1}(x_0)$. The isomorphism class of a connected covering is determined by the conjugacy class of its characteristic subgroup.*

4.5 Examples of covering spaces

Example 4.5.1 (Covers of \mathbb{S}^1 and \mathbb{C}^*). The example that illustrates all the basic features of this theory is $M = \mathbb{S}^1$ (which we will represent by unit complex numbers). The universal cover is clearly $\tilde{M} = \mathbb{R}$ with $\tilde{\pi}(x) = e^{2\pi i x}$. The deck transformations of \tilde{M} consist of integer shifts $x \mapsto x + n$, $n \in \mathbb{Z}$, therefore $\text{Aut}(\tilde{M}) = \mathbb{Z}$. And indeed, $\pi_1(\mathbb{S}^1) = \mathbb{Z}$. This universal cover can be interpreted as a slice of the Riemann surface of the complex logarithm $w = \ln z$, since the projection on this Riemann surface is exactly e^w . In fact, the entire Riemann surface (homeomorphic to \mathbb{C}) itself is the universal cover of $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Since \mathbb{C}^* and \mathbb{S}^1 are homotopy equivalent, their covering spaces are in one-to-one correspondence.

Furthermore, all subgroups of \mathbb{Z} have the form $k \cdot \mathbb{Z}$, $k \in \mathbb{N}$. The case $k = 0$ gives rise to the simply-connected universal cover with $\pi_1(\tilde{M}) = \{0\}$. Other values of k give rise to total spaces that are homeomorphic to \mathbb{S}^1 but the covering map is $\pi_k(e^{it}) = e^{ikt}$. These are the k -sheeted covers that can also be interpreted as slices of the Riemann surfaces of the complex roots $z^{1/k}$, since the projections on them are exactly z^k .

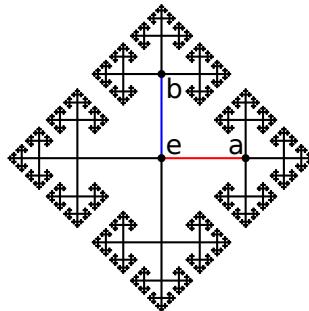


Fig. 4.3: This graph is the universal cover of the figure eight, which is $\Gamma = \mathbb{S}^1 \vee \mathbb{S}^1$. It is also the Cayley graph of the fundamental group of Γ , i.e. the free group with two generators a and b , $\pi_1(\Gamma) = F(\{a, b\}) \cong \mathbb{Z} * \mathbb{Z}$.

Example 4.5.2 (Universal covers of graphs). Let $M = \Gamma$ be a connected graph with V vertices and E edges, i.e. a topological space consisting of E copies of the interval $[0, 1]$ some of which are glued to each other at their ends. Such a graph always has a maximal tree, i.e. a subgraph of Γ that is a tree and is not contained in any other sub-tree of Γ . One can show that a maximal tree must contain all vertices of Γ . Since every tree is a contractible topological space, by retracting a maximal tree of Γ into one point we identify all vertices of Γ into one point. This shows that Γ is homotopy equivalent to the wedge sum of some number of circles, each corresponding to a cycle in the graph. Namely, the number l of cycles in Γ equals the number of edges in Γ minus the number of edges in a maximal tree of Γ . This is also equal to $1 - \chi(\Gamma)$, where $\chi(\Gamma) = V - E$ is the Euler characteristic. Then by Proposition 2.5.18, $\pi_1(\Gamma)$ is a free group with l generators. Finally, from the above results it is easy to see that $\tilde{\Gamma}$ is the **Cayley graph** of this free group, i.e. the graph whose vertices are all the elements of the group, and the edges incident to a vertex correspond to the generators of the group and their inverses. The vertex on the other end of an edge is the result of the multiplication of the current vertex by the edge's generator (say, from the right). In the case of a free group with l generators, this is a regular tree whose root represents the identity element and each vertex has valency/degree (number of incident edges) equal to l . See Fig. 4.3 for an example of a Cayley graph.

We can now use the theory of covering spaces to prove a very important basic theorem about free groups.

Theorem 4.5.3 (Nielsen-Schreier). *Any subgroup of a free group is free.*

Proof. Let G be a free group. As we've seen, it can be realized as the fundamental group

of a *bouquet* $M = \bigvee \mathbb{S}^1$ of as many circles as there are generators in G . Given a subgroup $H \leq G$, there must exist a covering space $E \xrightarrow{\pi} M$ such that $H = \pi_*(\pi_1(E, p_0))$ and $\pi_* : \pi_1(E, p_0) \rightarrow H \subset \pi_1(M, x_0)$ is an isomorphism. However, any covering space of a graph is a graph (indeed, each edge incident with a vertex in the base can be uniquely lifted to the total space, these liftings are homeomorphisms, and doing so for all vertices gives the total space a structure of a graph). In particular, any covering space of a bouquet of circles is a bouquet of circles. This implies that the fundamental group of E is free, and therefore H is free. ■

Corollary 4.5.4. *If a subgroup of a free group of finite rank has finite index, then it is also of finite rank.*

Proof. Indeed, if $H \leq G$ has finite index, it can be realized as the fundamental group of a finite covering graph of a finite bouquet of circles. But any such graph has a finitely generated fundamental group. ■

Theorem 4.5.5 (Schreier formula). *If G is a free group of finite rank r_G and $H \leq G$ is a subgroup of finite index n , then H is free of rank $r_H = n(r_G - 1) + 1$.*

Proof. Let $G = F(x_1, \dots, x_{r_G})$. This follows from examining the covering graph constructed in Theorem 4.5.3. It has one vertex for each right coset Hg and exactly $2r_G$ edges at each vertex (r_G going in and r_G going out) with the outgoing edge going from Hg to $Hgx_i = (Hg) \cdot x_i$ labeled by the generator x_i . Therefore this graph has $V = n$ vertices and $E = nr_G$ edges. The fundamental group of this graph $\tilde{\Gamma}$ is isomorphic to H and is generated by $1 - \chi(\tilde{\Gamma}) = 1 - V + E = n(r_G - 1) + 1$ loops. ■

Corollary 4.5.6. *Let Γ be a covering graph of $\bigvee_{i=1}^r \mathbb{S}^1$ (with $r > 1$) that corresponds to a subgroup generated by $r' > 1$ generators. If $n = (r' - 1)/(r - 1)$ is a positive integer then Γ has n sheets, otherwise Γ has infinitely many sheets.*

Corollary 4.5.7. *The free group of two generators, $F(a, b)$, contains free subgroups of any integer rank r_H .*

Proof. The set of all words whose degree (the sum of the powers of all generators in it) equals r_H is one such subgroup.

Alternatively, we can simply notice that the space $\mathbb{S}^1 \vee \mathbb{S}^1$ has covering spaces that are homeomorphic to $\bigvee_{i=1}^r \mathbb{S}^1$ for any r . ■

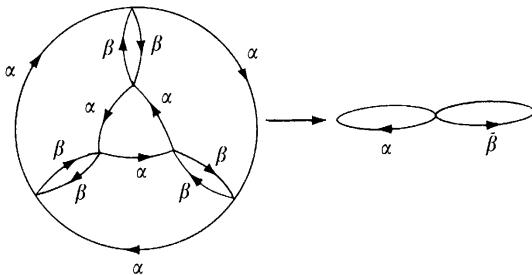


Fig. 4.4: A sixfold regular covering space of the figure eight. The fundamental group $\pi_1(E, p_0)$ is clearly generated by six loops. This makes it easy to guess that the stabilizer of any vertex is the normal subgroup of the free group $F(\alpha, \beta)$ generated by the six elements $\alpha^3, \beta^2, (\alpha\beta^{-1})^2, (\alpha^{-1}\beta)^2, \alpha\beta^2\alpha^{-1}, \alpha^{-1}\beta^2\alpha$. To find the group of deck transformations, we factor the free group by this normal subgroup, which is equivalent to adding these six generators as relations. It is easy to check that these relations define the permutation group $S_3 = \langle a, b \mid a^2, b^2, ababab \rangle$. If we reversed the arrows on the inner triangle of α 's, we'd get $\mathbb{Z}_2 \times \mathbb{Z}_3$ instead. From [19, Fig. III-7].

Example 4.5.8 (Covering spaces of $\mathbb{S}^1 \vee \mathbb{S}^1$). Consider the “figure eight” space $M = \mathbb{S}^1 \vee \mathbb{S}^1$. Its covering spaces are in one-to-one correspondence with conjugacy classes of subgroups of the free group $F(a, b)$ of two generators (one for each circle in M). By Theorem 4.5.3, any subgroup of a free group is free, so each covering space of M is uniquely determined by a subset of $F(a, b)$ that contains no relations (e.g. the set $\{a^2, b^2, a^2b^2\}$ is redundant and should be replaced with $\{a^2, b^2\}$). The elements of this subset generate the image of $\pi_1(E, p_0)$ under the monomorphism π_* .

Some examples of these covering spaces are given in Figures 4.4 and 4.5. The covering in Figure 4.2 is another one. It is instructive to observe which of the coverings in Figure 4.5 are regular, and how the subgroups corresponding to them are normal. It is similarly instructive to verify the Lifting Theorem 4.1.18: two different covering spaces of M are covering spaces of each other (say with basepoints) only if the generators of the stabilizer of one of them can be expressed in terms of the generators of the other. For example, (11) is a covering space of (1) because $b^nab^{-n} = (b^2)^{n/2}a(b^{-2})^{n/2}$ if n is even and $b^nab^{-n} = (b^2)^{(n-1)/2}bab^{-1}(b^{-2})^{(n-1)/2}$ otherwise.

(1) and (2), up to swapping a and b , are all of the existing connected double covers (so there are three in total). (3) and (4) are the same covering space but with different basepoints, and one can see how the corresponding subgroups are conjugates of each other by an element that connects those two points, for example b .

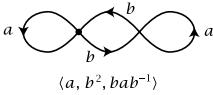
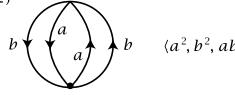
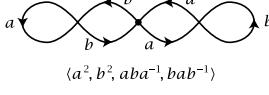
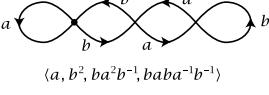
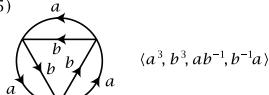
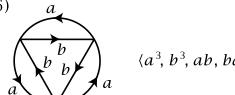
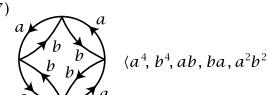
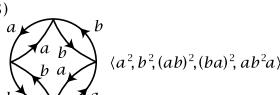
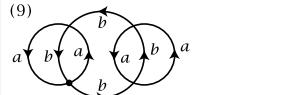
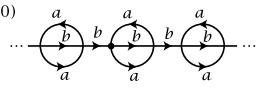
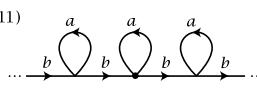
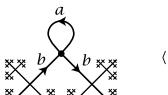
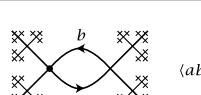
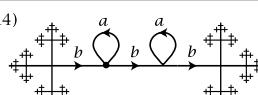
Some Covering Spaces of $S^1 \vee S^1$	
(1)	 $\langle a, b^2, bab^{-1} \rangle$
(2)	 $\langle a^2, b^2, ab \rangle$
(3)	 $\langle a^2, b^2, aba^{-1}, bab^{-1} \rangle$
(4)	 $\langle a, b^2, ba^2b^{-1}, baba^{-1}b^{-1} \rangle$
(5)	 $\langle a^3, b^3, ab^{-1}, b^{-1}a \rangle$
(6)	 $\langle a^3, b^3, ab, ba \rangle$
(7)	 $\langle a^4, b^4, ab, ba, a^2b^2 \rangle$
(8)	 $\langle a^2, b^2, (ab)^2, (ba)^2, ab^2a \rangle$
(9)	 $\langle a^2, b^4, ab, ba^2b^{-1}, bab^{-2} \rangle$
(10)	 $\langle b^{2n}ab^{-2n-1}, b^{2n+1}ab^{-2n} n \in \mathbb{Z} \rangle$
(11)	 $\langle b^nab^{-n} n \in \mathbb{Z} \rangle$
(12)	 $\langle a \rangle$
(13)	 $\langle ab \rangle$
(14)	 $\langle a, bab^{-1} \rangle$

Fig. 4.5: Examples of covering spaces of the figure eight space. Each one is a graph with a marked basepoint p_0 such that the neighborhood of each vertex locally looks the same as in the base space: two incoming edges (labeled a, b) and two outgoing (also labeled a, b). The label of an edge indicates which generator of $\pi_1(M)$ it is projected onto by π . Next to the space is the list of the generators of the local fundamental group $\pi_1(E, p_0)$ expressed in terms of a, b , which means that, viewed as part of $\pi_1(M)$, they generate the stabilizer $G_{p_0} = \pi_*(\pi_1(E, p_0))$. The number of sheets matches the index of this subgroup. Two vertices are related by a deck transformation iff their stabilizers coincide, otherwise the stabilizers are merely conjugate. For regular coverings (1,2,5,6,7,8,9,11), all stabilizers coincide and therefore are normal. From [22, p. 58].

Exercise 4.1. Show that the covering space of the figure eight space $M = \mathbb{S}^1 \vee \mathbb{S}^1$ that corresponds to the normal subgroup generated by a^2 , b^2 , and $(ab)^4$ (where a, b are the two generators of the free group $\pi_1(M)$) is regular, and describe what it looks like. *Hint:* this subgroup has index 8, so it's an eight-fold covering.

Exercise 4.2. Describe the covering space of the figure eight space that corresponds to the commutator subgroup of the free group of two generators. *Hint:* $F(a, b)/[F(a, b), F(a, b)] = \mathbb{Z}^2$, so this covering space has a \mathbb{Z}^2 -symmetry.

Exercise 4.3. Identify in \mathbb{S}^1 the open upper and the open lower semicircle to a point. The resulting space X has four points. Show $\pi_1(X) \cong \mathbb{Z}$. Does X have a universal covering?

Exercise 4.4. Show that the universal cover of $M_1 \times M_2$ is $\tilde{M}_1 \times \tilde{M}_2$.

Exercise 4.5 ([22, Exercise B.7]). If G is a finitely generated free group and N is a nontrivial normal subgroup of infinite index, show, using covering spaces, that N is not finitely generated.

Example 4.5.9. $\widetilde{\mathbb{T}^n} = \mathbb{R}^n$, $\widetilde{\mathrm{SO}_3} = \mathrm{SU}_2 = \mathbb{S}^3$, $\widetilde{\mathrm{SO}_{n>2}} = \mathrm{Spin}_n$, $\widetilde{\mathrm{U}_n} = \mathrm{SU}_n \times \mathbb{R}$.

5 Homotopy Theory

5.1 Homotopy groups

Definition 5.1.1 (Homotopy groups). Pick the base points $\bullet \in \mathbb{S}^n$ and $x_0 \in X$. The set of homotopy classes of maps $\gamma : \mathbb{S}^n \rightarrow X$ such that $\gamma(\bullet) = x_0$ forms the n -th homotopy group $\pi_n(X, x_0) = [\mathbb{S}^n; X]_\bullet$. For two such maps, define their product by first mapping \mathbb{S}^n onto the wedge sum $\mathbb{S}^n \vee \mathbb{S}^n$ via contracting an arbitrarily chosen equator (which is homeomorphic to \mathbb{S}^{n-1}) into one point, then apply γ_1 and γ_2 to the two hemispheres respectively, making the equator point the base point for them.

For $n = 0$, $\pi_0(X)$ is defined as the set of path-connected components of X (which is also the set of homotopy classes of pointed maps $\mathbb{S}^0 \rightarrow X$) but does not have a natural group structure.

Exercise 5.1. 1. The above definition is consistent and the resulting homotopy classes don't in fact depend on the choice of an equator. In fact, in natural terms, this group operation is exactly the coproduct on the category of pointed spaces, given by the wedge sum $\gamma_1 \vee \gamma_2 : \mathbb{S}^n \vee \mathbb{S}^n \rightarrow X$ (a.k.a. *concatenation*).
2. All homotopy groups for different x_0 are isomorphic.

Proposition 5.1.2 (Functoriality of π_n). Any pointed continuous map $f : X \rightarrow Y$ induces a homomorphism of the homotopy groups $f_* : \pi_n(X, x_0) \rightarrow \pi_n(Y, y_0)$, and $(g \circ f)_* = g_* \circ f_*$. In other words, we have a functor $\pi_n : \text{Top}_\bullet \rightarrow \text{Gr}$.

Theorem 5.1.3. Homotopy groups π_n , $n \geq 1$, are homotopy invariants, i.e. if $X \simeq Y$ are two homotopy equivalent path connected spaces, then $\pi_n(X, x_0) \cong \pi_n(Y, y_0)$.

Proof. Exercise. Also see [22]. ■

Corollary 5.1.4. π_n descends to a functor $\pi_n : \text{hTop}_\bullet \rightarrow \text{Gr}$ on the category hTop_\bullet , in which the morphisms are the homotopy classes of pointed continuous maps modulo homotopy relative to the base point. In particular, homotopy equivalent spaces have isomorphic homotopy groups.

Theorem 5.1.5. For $n \geq 2$, all groups $\pi_n(X, x_0)$ are abelian.

Proof. Follows from Proposition 2.5.14 and the fact that $\mathbb{S}^n = \Sigma \Sigma \mathbb{S}^{n-2}$ for $n \geq 2$. We have $[\mathbb{S}^n, Y] = [\Sigma \Sigma \mathbb{S}^{n-2}, Y] = [\Sigma \mathbb{S}^{n-2}, \Omega Y]$, which is abelian. ■

Another very important connection between homotopy groups and suspension will be discussed when we prove Freudenthal's Suspension Theorem 5.11.8.

5.2 Exact sequence of homotopy groups

Definition 5.2.1 (Topological triple). A triple (X, A, B) is a topological space X with a pair of subspaces $A \subset B \subset X$. When $B = \bullet$ is a point, the triple is also called a pointed pair.

Definition 5.2.2 (Relative homotopy groups). For a pointed pair (X, A, \bullet) and $n \geq 1$ the relative homotopy group $\pi_n(X, A, \bullet)$ is the set of homotopy classes of triples $\pi_n(X, A, \bullet) = [(\mathbb{D}^n, \mathbb{S}^{n-1}, \bullet); (X, A, \bullet)]$, where \mathbb{D}^n is the n -disk, \mathbb{S}^{n-1} is its boundary, and the homotopies are through pointed maps that take \mathbb{S}^{n-1} into A . Note that this is *not* the same as homotopies relative to \mathbb{S}^{n-1} .

These are groups with the same multiplication operation and they are still abelian for $n \geq 3$ because $[(\Sigma X, \Sigma A); (Y, B)] = [(X, A); (\Omega Y, \Omega B)]$.

Proposition 5.2.3 (Compression criterion). A map $f : (\mathbb{D}^n, \mathbb{S}^{n-1}, \bullet) \rightarrow (X, A, \bullet)$ represents the identity in $\pi_n(X, A, \bullet)$ iff it is homotopic rel $\partial\mathbb{D}^n \cong \mathbb{S}^{n-1}$ to a map whose image is contained in A .

Proof. For necessity, a map whose image is contained in A can be composed with a deformation retraction of \mathbb{D}^n onto the basepoint to give a homotopy to a trivial map. Conversely, if $[f] = e$ via a homotopy $H : \mathbb{D}^n \times I \rightarrow X$, then by restricting H to a family of n -disks in $\mathbb{D}^n \times I$ starting with $\mathbb{D}^n \times \{0\}$ and ending with the disk $\mathbb{D}^n \times \{1\} \cup \mathbb{S}^{n-1} \times I$, all the disks in the family having the same boundary, then we get a homotopy from f to a map onto A , stationary on \mathbb{S}^{n-1} . This is a homotopy relative to A . ■

Any map $f : (X, A, x_0) \rightarrow (Y, B, y_0)$ induces maps $f_* : \pi_n(X, A, x_0) \rightarrow \pi_n(Y, B, y_0)$ which are homomorphisms for $n \geq 2$ and have properties analogous to those in regular homotopy: $(f \circ g)_* = f_* \circ g_*$, $\text{id}_* = \text{id}$, and $f_* = g_*$ if $f \sim g$ through maps of pointed pairs $(X, A, x_0) \rightarrow (Y, B, y_0)$.

In particular, the inclusions $i : (A, x_0) \hookrightarrow (X, x_0)$ and $j : (X, x_0, x_0) \hookrightarrow (X, A, x_0)$ induce such homomorphisms. Moreover, there is a third special homomorphisms that connects homotopy groups in adjacent degrees.

Definition 5.2.4 (Boundary operator for homotopy). The inclusion of the boundary $\mathbb{S}^{n-1} \hookrightarrow \mathbb{D}^n$ induces restrictions of maps $f : (\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, A, x_0)$ to the boundary, i.e. $\partial f = f|_{\partial\mathbb{D}^n}$. For $n \geq 2$ this operator defines a homomorphism of relative homotopy groups.

Definition 5.2.5 (Exact sequence). Let C be a category in which the notions of kernels and images of morphisms are well-defined. For now we assume that it is a concrete category, so kernels and images are understood as subsets of the domain and codomain, respectively.

Next let there be a (finite or infinite) sequence of morphisms

$$\dots \rightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \rightarrow \dots \quad (5.2.1)$$

We say that this sequence is exact at Y if $\text{im } f = \ker g$.

Example 5.2.6. In the category of pointed sets (or topological spaces etc.), kernels can be defined as the pre-images of the basepoint. Exactness then amounts to $\text{im } f = g^{-1}(\ast)$. Usually we will be concerned with the case $C = \text{Gr}$.

Definition 5.2.7 (Short exact sequence). A short exact sequence is a sequence of the form

$$0 \rightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \rightarrow 0, \quad (5.2.2)$$

where 0 is a zero object in C .

The exactness at X is equivalent to f being a monomorphism. Exactness at Z is equivalent to g being an epimorphism. The exactness at Y essentially says that $Z \cong Y/\text{im } f = Y/\ker g \cong Y/X$. In the category of groups this is just the first isomorphism theorem.

Example 5.2.8. If the short exact sequence is even shorter, e.g.

$$0 \rightarrow X \xrightarrow{f} Y \rightarrow 0, \quad (5.2.3)$$

then its exactness is equivalent to f being an isomorphism.

Theorem 5.2.9 (Long exact sequence of homotopy groups). *The sequence of homomorphisms*

$$\dots \rightarrow \pi_n(A, x_0) \xrightarrow{i_*} \pi_n(X, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0) \xrightarrow{\partial} \pi_{n-1}(A, x_0) \rightarrow \dots \rightarrow \pi_0(X, x_0) \quad (5.2.4)$$

is exact. Near the end of the sequence, where the group structure is not defined, the “kernel” is the path-connected component of x_0 .

Proof. In fact we will prove the long exact sequence of a pointed triple (X, A, B, x_0) with $x_0 \in B \subset A \subset X$:

$$\dots \rightarrow \pi_n(A, B, x_0) \xrightarrow{i_*} \pi_n(X, B, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0) \xrightarrow{\partial} \pi_{n-1}(A, B, x_0) \rightarrow \dots \rightarrow \pi_1(X, A, x_0). \quad (5.2.5)$$

The boundary operator for triples is defined as the composition $\partial : \pi_n(X, A) \xrightarrow{\partial} \pi_{n-1}(A) \xrightarrow{i_*} \pi_{n-1}(A, B)$ where $i : B \hookrightarrow A$ is the inclusion map. When $B = \{x_0\}$ this reduces to the exact

sequence of the pointed pair (X, A, x_0) , though the latter sequence continues for two more steps to $\pi_0(X, x_0)$.

Exactness at $\pi_n(X, B, x_0)$: First note that the composition $j_* \circ i_*$ is trivial since every map $(\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (A, B, x_0)$ represents zero in $\pi_n(X, A, x_0)$ by the compression criterion. To see that $\ker j_* \subset \text{im } i_*$, let $f : (\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, B, x_0)$ represent the trivial element in $\pi_n(X, A, x_0)$. Then by the compression criterion again, f is homotopic rel \mathbb{S}^{n-1} to a map with image in A , hence $[f] \in \pi_n(X, B, x_0)$ is in the image of i_* .

Exactness at $\pi_n(X, A, x_0)$: The composition $\partial \circ j_*$ is zero since the restriction of a map $(\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, B, x_0)$ to \mathbb{S}^{n-1} has image lying in B , and hence represents the trivial class in $\pi_{n-1}(A, B, x_0)$. Conversely, suppose the restriction of $f : (\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, A, x_0)$ to \mathbb{S}^{n-1} represents the trivial element $\pi_{n-1}(A, B, x_0)$. Then $f|_{\mathbb{S}^{n-1}}$ is homotopic to a map with image in B via a pointed homotopy $H : \mathbb{S}^{n-1} \times I \rightarrow A$. We can tack H onto f to get a new map $(\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, B, x_0)$ which, as a map $(\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, A, x_0)$, is homotopic to f by the homotopy that tacks on increasingly longer initial segments of H . So $[f] \in \text{im } j_*$.

Exactness at $\pi_n(A, B, x_0)$: The composition $i_* \circ \partial$ is trivial since the restriction of a map $f : (\mathbb{D}^{n+1}, \mathbb{S}^n, s_0) \rightarrow (X, A, x_0)$ to \mathbb{S}^n is homotopic rel s_0 to a constant map via f itself (via a contraction of \mathbb{S}^n onto s_0 inside \mathbb{D}^{n+1}). The converse is easy if B is a point, since the nullhomotopy $f_t : (\mathbb{D}^n, \mathbb{S}^{n-1}) \rightarrow (X, x_0)$ of $f_0 : (\mathbb{D}^n, \mathbb{S}^{n-1}) \rightarrow (A, x_0)$ gives a map $F : (\mathbb{D}^{n+1}, \mathbb{S}^n, s_0) \rightarrow (X, A, x_0)$ with $\partial([F]) = [f_0]$. Thus the proof is finished in this case.

For a general B , let $f : (\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (A, B, x_0)$ be in the kernel of i_* and let $F : \mathbb{D}^n \times I \rightarrow X$ be a corresponding nullhomotopy of f through maps $(\mathbb{D}^n, \mathbb{S}^{n-1}, s_0) \rightarrow (X, B, x_0)$. It is then clear that F can be viewed as a map $\mathbb{D}^{n+1} \rightarrow X$ which restricts to f on $\mathbb{D}^n \times \{0\}$, and maps the rest of the boundary into $B \subset A$, which makes f the boundary of some map $(\mathbb{D}^{n+1}, \mathbb{S}^n, s_0) \rightarrow (X, A, x_0)$. ■

5.3 CW-complexes

Definition 5.3.1 (CW-complex). A cellular structure on a set X is a family \mathcal{F} of *characteristic mappings* $f_\alpha^n : \mathbb{D}^n \rightarrow X$, where $n = 0, 1, 2, \dots$ and $\alpha \in A_n$ (some indexing sets), such that the following conditions hold. Let $X^{(n)}$ denote the union of images of mappings f_α^k , $k \leq n$. Also $X^{(-1)} = \emptyset$.

1. For every n , $\bigsqcup_\alpha f_\alpha^n$ maps $\bigsqcup_\alpha \text{int } \mathbb{D}^n$ injectively to $X \setminus X^{(n-1)}$. Here, we put $\text{int } \mathbb{D}^0 = \mathbb{D}^0$.
2. Every f_α^n maps $\partial \mathbb{D}^n$ to $X^{(n-1)}$.
3. $X = \bigcup_n X^{(n)}$.

A CW-complex is a Hausdorff topological space with a CW-structure such that the topology coincides with the final topology defined by \mathcal{F} , also called the *weak topology*: $A \subset X$ is open iff $A \cap X^{(n)}$ is open in $X^{(n)}$ for each n .

The restrictions $\partial f_\alpha^n = f_\alpha^n|_{\partial \mathbb{D}^n}$ are called the *attaching maps*. The subsets $X^{(n)} \subset X$ are called

the n -skeleta of \mathcal{F} . The images $f_\alpha^n(\mathbb{D}^n)$ are referred to as the *closed cells* and $f_\alpha^n(\text{int } \mathbb{D}^n)$ as *open cells* (even though these are actually closed or open only as subspaces of $X^{(n)}$!). A CW-complex (X, \mathcal{F}) is said to be pointed if X is pointed and the basepoint is a 0-cell. A subcomplex is a subspace $X' \subset X$ endowed with the relative topology, together with a subfamily $\mathcal{F}' \subset \mathcal{F}$ such that (X', \mathcal{F}') is a CW-complex. The highest dimension of a cell in a complex is called its dimension. A CW-complex is called finite if it is finite-dimensional and has a finite number of cells.

We will typically work with finite CW-complexes, but infinite-dimensional ones will be useful as well.

Remark 5.3.2. The acronym CW refers to the following properties.

1. Closure-finiteness: every closed cell meets only finitely many open cells, which we will prove in Proposition 5.3.5.
2. Weak topology: X carries the final topology defined by the family \mathcal{F} .

A more general concept of cellular complexes drops these requirements. Note that, due to the inductive definition of CW-complexes as a collection of attaching maps, we get (C) automatically.

Every CW-complex therefore can be thought of as a *decomposition* of a Hausdorff space X into a disjoint union of open cells of various dimensions. The space X , up to homeomorphism, can be inductively reconstructed from its CW-structure as follows. First, $X^{(0)}$ is a discrete set of points. Next, given an $(n-1)$ -skeleton $X^{(n-1)}$, we can obtain $X^{(n)}$ by taking a disjoint union of n -disks $\sqcup_\alpha \mathbb{D}^n$ and identifying their boundaries with points of $X^{(n-1)}$ via the attaching maps $\partial f_\alpha^n : \partial \mathbb{D}^n \rightarrow X^{(n-1)}$, which are required to be continuous. The space X is then the inductive union of all skeleta with the final topology.

Example 5.3.3.

1. A graph is nothing but a 0- or 1-dimensional cell complex.
2. The infinite bouquet $\bigvee_{i=1}^\infty \mathbb{S}^1$ is not a CW-complex because the basepoint has to meet infinitely many 1-cells.
3. The sphere \mathbb{S}^n admits a CW-structure with one 0-cell and one n -cell. The characteristic maps can be chosen as

$$f^0(*) = e_1, \quad f^n(x) = (2|x|^2 - 1, 2\sqrt{1 - |x|^2}x). \quad (5.3.1)$$

Another useful CW-structure on the sphere has two cells in each dimension up to n . Its characteristic maps are

$$f_\pm^0(*) = \pm e_1 \quad f_\pm^k(x) = (x, \pm\sqrt{1 - |x|^2}, 0, \dots, 0). \quad (5.3.2)$$

The two k -cells are given by

$$\{(x_1, \dots, x_{k+1}, 0, \dots, 0) \in \mathbb{S}^n : \pm x_{k+1} \geq 0\}. \quad (5.3.3)$$

4. The closed n -disk \mathbb{D}^n has a tautological CW-structure with one n -cell. It is however sometimes convenient to have the boundary \mathbb{S}^{n-1} as a subcomplex. This can be achieved by just adding either one of the two CW-structures of \mathbb{S}^{n-1} above.
5. The wedge sum of two CW-complexes $(X_1, \mathcal{F}_1) \vee (X_2, \mathcal{F}_2)$ carries a natural CW-structure $\mathcal{F}_1 \cup \mathcal{F}_2$.
6. The direct product $(X_1, \mathcal{F}_1) \times (X_2, \mathcal{F}_2)$ has a CW-structure with elements $(f_{1i}^n \times f_{2j}^m) \circ p_{n+m}$ where $p_{n+m} : \mathbb{D}^{n+m} \rightarrow \mathbb{D}^n \times \mathbb{D}^m$ is some chosen homeomorphism. This makes sense, because, as a homeomorphism, p_{n+m} maps the boundary \mathbb{S}^{n+m-1} of \mathbb{D}^{n+m} onto the boundary $(\mathbb{S}^{n-1} \times \mathbb{D}^m) \cup (\mathbb{D}^n \times \mathbb{S}^{m-1})$ of $\mathbb{D}^n \times \mathbb{D}^m$. For example, the direct product of two copies of \mathbb{S}^1 with one cell in dimensions 0 and 1 yields a CW-structure on the torus \mathbb{T}^2 with one 0-cell, one 2-cell, and two 1-cells. Note that in general the topology on the product complex need not coincide with the product topology, but it does when the complexes are locally finite (i.e. every point has a neighborhood that overlaps only with finitely many open cells – such spaces are also paracompact).

Proposition 5.3.4 ([17, Prop. 3.1.11]). *Let (X, \mathcal{F}) be a CW-complex, Y a topological space, and $f : X \rightarrow Y$ a map. Then f is continuous iff its restrictions $f \circ f_\alpha^n$ to each cell are continuous.*

Proof. This is nothing but the universal property of the final topology, but let us walk through the argument. Only the sufficiency is not obvious. Let $V \subset Y$ be open. By assumption, $(f \circ f_\alpha^n)^{-1}(V)$ is open in \mathbb{D}^n for all n and α . Since $(f \circ f_\alpha^n)^{-1}(V) = (f_\alpha^n)^{-1}(f^{-1}(V))$, then $f^{-1}(V) \subset X$ is open by virtue of being the image of an open set under a quotient map (and all quotient maps are open by definition). ■

Proposition 5.3.5 ([19, Prop. 8.1]). *If X is a CW-complex then*

1. *if $A \subset X$ has no two points in the same open cell, then A is closed and discrete;*
2. *if $K \subset X$ is compact then K is contained in a finite union of open cells;*
3. *each cell of X is contained in a finite subcomplex of X (i.e. the cell complex is “closure finite”).*

Proof. We will first show that all three statements are equivalent to each other.

(1) \Rightarrow (2). If $K \subset X$ is compact then let $A \subset K$ be a set of points, one from each open cell touching K . By (1) A is closed, hence compact, and discrete. Hence A is finite. But that means K is contained in a finite union of open cells as claimed.

(2) \Rightarrow (3). In fact, for a cell α , we will only use (2) for the set K which is the image of the attaching map $f_{\partial\alpha} : \partial\mathbb{D}^n \rightarrow X^{(n-1)}$ of that cell, and the images of attaching maps of smaller-dimensional cells. Statement (2) implies that $X_\sigma = f_\sigma(\mathbb{D}^n)$ is contained in a finite union of open cells, and by construction all of these are of smaller dimension except for

$U_\sigma = f_\sigma(\text{int } \mathbb{D}^n)$ itself. By the same token each of these lower-dimensional closed cells is contained in a finite union of open cells of even lower dimension (except for open the cell itself), and so on. This reasoning obviously must come to an end with 0-cells in a finite number of steps, and, at that stage, the union of the cells produced a finite subcomplex.

(3) \Rightarrow (1). Consider the intersection of A with a closed cell. By (3) this is contained in a finite subcomplex. Since A has at most one point in common with any open cell this intersection must be finite, and hence closed. For any point $x \in A$, the set $A \setminus \{x\}$ satisfies the hypothesis for A , and so it must also be closed. Hence $\{x\}$ is open in A , so A is discrete. Finally, we put all this together. Statement (1) clearly holds for $X^{(0)}$ (it must be discrete because it has the final topology of maps of points). Suppose we know (1) for the n -skeleton $X^{(n)}$. Then we also have (2) for $X^{(n)}$. In turn, we get (3) for $X^{(n)}$. But the proof of (2) \Rightarrow (3) for a particular k -cell only used (2) for subsets of $X^{(k-1)}$. Thus we actually have (3) for $X^{(n+1)}$. Thus we also have (1) for $X^{(n+1)}$. Consequently, we have all three statements for $X^{(n)}$ for all n . But any cell is in some $X^{(n)}$, and so we know (3) for X itself, and this implies (1) then (2) for X . ■

Corollary 5.3.6. *The boundary of each cell is contained in finitely many closed cells of lower dimensions.*

Theorem 5.3.7 ([19, Thm. 8.2]). *A compact subset of a CW-complex is contained in a finite subcomplex.*

Proof. Let $K \subset X$ be compact. By (2) of Proposition 5.3.5, K is contained in a union of a finite number of open cells. By (3) of the same Proposition, each of these is contained in a finite subcomplex. The union of this finite number of finite subcomplexes is a finite subcomplex containing K . ■

Corollary 5.3.8. *If $X_1 \hookrightarrow X_2 \hookrightarrow \dots$ is an infinite sequence of inclusions of CW-subcomplexes, then*

$$\pi_n(\varinjlim X_i) = \varinjlim \pi_n(X_i). \quad (5.3.4)$$

Proof. Let $X = \varinjlim X_i = \bigcup X_i$ be the limiting CW-complex. Since \mathbb{S}^n is compact, its continuous image in X under any representative of an element of $\pi_n(X)$ is compact and thus contained in a finite subcomplex. Therefore every element of $\pi_n(X)$ is contained in $\pi_n(X_k)$ for some k and hence in the directed colimit $\varinjlim \pi_n(X_i)$. Conversely, every element of this colimit is represented by some element of $\pi_n(\overrightarrow{X_k})$ for some k by the explicit construction of the colimit. ■

Proposition 5.3.9. A CW-complex is locally compact (every point has an open neighborhood contained in a compact set) iff it is locally finite (every point has an open neighborhood contained in a finite union of cells).

Proof. Exercise. ■

Definition 5.3.10 (Directed system, Direct limit). A directed system of topological spaces consists of a directed set (A, \leq) (a poset such that for all $a, b \in A$ there exists $c \in A$ such that $a \leq c$ and $b \leq c$), a topological space X_α for every $\alpha \in A$ and a continuous mapping $f_{\beta\alpha} : X_\alpha \rightarrow X_\beta$ for every pair $\alpha \leq \beta$ such that $f_{\alpha\alpha} = \text{id}_{X_\alpha}$ and $f_{\gamma\beta} \circ f_{\beta\alpha} = f_{\gamma\alpha}$ for all $\alpha \leq \beta \leq \gamma$. The direct limit (*colimit*)

$$X = \varinjlim X_\alpha \quad (5.3.5)$$

of the directed system $\{X_\alpha, f_{\beta\alpha}\}$ is (up to homeomorphism) the topological quotient of the disjoint union $\bigsqcup_{\alpha \in A} X_\alpha$ with respect to the equivalence relation that $x \in X_\alpha$ is equivalent to $y \in X_\beta$ iff $\gamma \in A : f_{\gamma\alpha}(x) = f_{\gamma\beta}(y)$. Composition of the natural inclusions $X_\alpha \hookrightarrow \bigsqcup_\alpha X_\alpha$ with the quotient map yields continuous mappings

$$\varphi_\alpha : X_\alpha \rightarrow X \quad (5.3.6)$$

and the topology of X coincides with the final topology defined by these maps. That is, a subset of X is open iff its pre-image under φ_α is open in X_α for all α . In other words, the direct limit space is the set-theoretic direct limit with the final topology for the maps φ_α .

Direct limits are nothing but a generalization of pushouts to infinite families of objects, and will come up again in a more general categorical setting when we discuss homological algebra.

Proposition 5.3.11 ([17, Prop. 3.1.14]). Let $\{X_\alpha, f_{\beta\alpha}\}$ and $\{Y_\alpha, g_{\beta\alpha}\}$ be directed systems of topological spaces over the same index set A and let X and Y , respectively, be the direct limits. Every family of continuous mappings $h_\alpha : X_\alpha \rightarrow Y_\alpha$ satisfying $h_\beta \circ f_{\beta\alpha} = g_{\beta\alpha} \circ h_\alpha$ whenever $\alpha \leq \beta$ descends to a continuous mapping $h : X \rightarrow Y$.

Proposition 5.3.12 ([17, Prop. 3.1.15]). Let $\{X_\alpha, f_{\beta\alpha}\}$ be a directed system of topological spaces and let X be the direct limit. If for some i one has $\pi_i(X_\alpha) = 0$ for almost all (i.e. all but finitely many) α , then $\pi_i(X) = 0$.

Theorem 5.3.13 (CW Construction Theorem [8, Thm. 5.20]). Suppose $X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots$ is a sequence of topological spaces satisfying the following conditions:

- (i) X_0 is a nonempty discrete space.

(ii) For each $n \geq 1$, X_n is obtained from X_{n-1} by attaching a (possibly empty) collection of n -cells.

Then $X = \bigcup_n X_n$ has a unique topology coherent with the family $\{X_n\}_n$ (i.e. it is the final topology for the inclusions $X_n \hookrightarrow X$), and a unique cell decomposition making it into a CW-complex whose n -skeleton is X_n for each n .

Proof. The uniqueness of the topology coherent with $\{X_n\}_n$ and of the cell decomposition is practically obvious. Finally, in the case of infinitely many cells one needs to check that X is closure-finite and has the weak topology. See [8, Thm. 5.20] for full details. ■

Corollary 5.3.14. *The family of skeleta $X^{(k)}$ of a CW-complex together with the natural inclusion maps $f_{lk} : X^{(k)} \hookrightarrow X^{(l)}$ for $k \leq l$ forms a directed system of spaces. The direct limit of this system is homeomorphic to X :*

$$X \cong \varinjlim X^{(n)}. \quad (5.3.7)$$

Example 5.3.15 (\mathbb{R}^∞). By including $\mathbb{R}^n \hookrightarrow \mathbb{R}^\infty$ via $x = (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0, 0, \dots)$, we can take the direct limit

$$\mathbb{R}^\infty = \varinjlim \mathbb{R}^n. \quad (5.3.8)$$

This space consists of sequences of reals with only finitely many nonzero elements. The final topology on this space is strictly finer than the subspace topology inherited on it from the direct product topology on \mathbb{R}^∞ .¹⁷ In fact, \mathbb{R}^∞ itself is not open as a subset of \mathbb{R}^∞ .

Example 5.3.16 (\mathbb{S}^∞). Since we can obtain \mathbb{S}^n from \mathbb{S}^{n-1} by attaching two cells, we can take the direct limit

$$\mathbb{S}^\infty = \varinjlim \mathbb{S}^n. \quad (5.3.9)$$

This space includes every sphere \mathbb{S}^n as a subcomplex. By the above theorem, $\pi_n(\mathbb{S}^\infty) = 0$ for all n , however let us show the stronger result that \mathbb{S}^∞ is contractible.

The space \mathbb{R}^∞ has a norm inherited from Euclidean norms on \mathbb{R}^n , which is simply $|x| = \sqrt{\sum_{i=1}^{\infty} x_i^2}$. The unit sphere under this norm is clearly a union of regular finite-dimensional spheres embedded into each other at the equator, which is \mathbb{S}^∞ . First, define the homotopy $f_t : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$ by

$$f_t(x) = (1 - t)x + tTx, \quad (5.3.10)$$

where $T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ is the shift operator. This f_t takes nonzero vectors to nonzero vectors for all $t \in [0, 1]$, therefore $f_t/|f_t|$ restricts to a homotopy from $\text{id}_{\mathbb{S}^\infty}$ to Tx on

\mathbb{S}^∞ . Then a homotopy from this map to a constant map is given by $g_t/|g_t|$ where

$$g_t(x) = (1 - t)Tx + te_1, \quad (5.3.11)$$

where $e_1 = (1, 0, 0, \dots)$ is the basepoint.

Example 5.3.17 (CW-structures of projective spaces). The projective space $\mathbb{R}P^n$ is defined as the space of all lines through the origin in \mathbb{R}^{n+1} , i.e. the topological quotient of $\mathbb{R}P^n \setminus \{0\}$ under the identifications $v \sim \lambda v$ for all $v \in \mathbb{R}^{n+1} \setminus \{0\}$ and $\lambda \neq 0$. We can restrict to vectors of length 1, so it becomes the quotient space $\mathbb{S}^n/(v \sim -v)$. This is equivalent to the quotient of a hemisphere \mathbb{D}^n with antipodal points of $\partial\mathbb{D}^n$ identified. Since $\partial\mathbb{D}^n$ with antipodal points identified is just $\mathbb{R}P^{n-1}$, we see that $\mathbb{R}P^n$ is obtained from $\mathbb{R}P^{n-1}$ by attaching an n -cell, with the quotient map $\mathbb{S}^{n-1} \rightarrow \mathbb{R}P^{n-1}$ as the attaching map. It follows by induction on n that $\mathbb{R}P^n$ has a cell complex structure $e^0 \cup e^1 \cup \dots \cup e^n$ with one cell e^i in each dimension $i \leq n$.

As a consequence, the infinite union $\mathbb{R}P^\infty = \bigcup_n \mathbb{R}P^n$ becomes a cell complex with one cell in each dimension. We can view $\mathbb{R}P^\infty$ as the space of lines through the origin in \mathbb{R}^∞ .

Similarly, it is possible to obtain $\mathbb{C}P^n$, the space of complex lines through the origin in \mathbb{C}^{n+1} , as the quotient of the disk \mathbb{D}^{2n} under the identifications $v \sim \lambda v$ for $v \in \partial\mathbb{D}^{2n}$ and $|\lambda| = 1$, in the following way. The vectors in the unit sphere $\mathbb{S}^{2n+1} \subset \mathbb{C}^{n+1}$ with last coordinate real and nonnegative are precisely the vectors of the form $(w, \sqrt{1 - |w|^2}) \in \mathbb{C}^n \times \mathbb{C}$ with $|w| \leq 1$. Such vectors form the graph of the function $w \mapsto \sqrt{1 - |w|^2}$. This is a disk \mathbb{D}_+^{2n} bounded by the sphere $\mathbb{S}^{2n-1} \subset \mathbb{S}^{2n+1}$ consisting of vectors $(w, 0) \in \mathbb{C}^n \times \mathbb{C}$ with $|w| = 1$. Each vector in \mathbb{S}^{2n+1} is equivalent under the identifications $v \sim \lambda v$ to a vector in \mathbb{D}_+^{2n} , and the latter vector is unique if its last coordinate is nonzero. If the last coordinate is zero, we have just the identifications $v \sim \lambda v$ for $v \in \mathbb{S}^{2n-1}$.

From the description of $\mathbb{C}P^n$ as the quotient of \mathbb{D}_+^{2n} under the identifications $v \sim \lambda v$ for $v \in \mathbb{S}^{2n-1}$ it follows that $\mathbb{C}P^n$ is obtained from $\mathbb{C}P^{n-1}$ by attaching a cell e^{2n} via the quotient map $\mathbb{S}^{2n-1} \rightarrow \mathbb{C}P^{n-1}$. So by induction on n we obtain a cell structure $\mathbb{C}P^n = e^0 \cup e^2 \cup \dots \cup e^{2n}$ with cells only in even dimensions. Similarly, $\mathbb{C}P^\infty$ has a cell structure with one cell in each dimension.

Theorem 5.3.18 ([8, Thm. 5.22]). *Every CW-complex is paracompact.*

Proof. It suffices to show that any open covering $\{U_\alpha\}_\alpha$ of the complex X supports a subordinate **POU** (since part of the definition of a **POU** is that only finitely many of the functions χ_α have non-zero values at a given point). It can be constructed inductively by extending a **POU** from the n -skeleton to the $(n+1)$ -skeleton, extending each function cell by cell.

A full proof can be found in [8, Thm. 5.22], but one needs to include the corrections from the book's **errata list**. ■

Proposition 5.3.19 ([8, Prop. 5.23]). Suppose X is a CW-complex with countably many cells. If X is locally Euclidean, then it is a topological manifold.

Proof. Every CW-complex is Hausdorff by definition. Further, X is a quotient of a disjoint union of countably many disks of various dimensions. Such a disjoint union is easily seen to be second countable. One then uses the general fact that quotient maps preserve second countability (it is enough to check the existence of a finite subcover for any cover by coordinate balls, because each coordinate ball is second countable and a union of countably many such sets has to be as well). We conclude that X is second countable, and thus a manifold. ■

5.4 Homotopy extension

Definition 5.4.1 (Homotopy Extension Property). A continuous map $i : A \rightarrow X$ is said to have the homotopy extension property with respect to a space Y if any homotopy $F : A \times I \rightarrow Y$ of a given map $F_0 : A \rightarrow Y$ can be extended to a homotopy $\tilde{F} : X \times I \rightarrow Y$ with any given initial condition \tilde{F}_0 . In other words, we are looking for a continuous map \tilde{F} that makes the following diagram commute:

$$\begin{array}{ccc} Y & \xleftarrow{\tilde{F}_0} & X \\ e_0 \uparrow & \swarrow \tilde{F} & \uparrow i \\ Y^I & \xleftarrow{F} & A. \end{array} \quad (5.4.1)$$

Here, e_0 is the evaluation map $\gamma \in Y^I \mapsto \gamma(0) \in Y$.

A topological pair (X, A) is said to have the **Homotopy Extension Property (HEP)** with respect to Y if the inclusion map $i : A \hookrightarrow X$ does.

Definition 5.4.2 (Cofibration). If a map $i : A \hookrightarrow X$ satisfies the **HEP** for any space Y , it is called a cofibration.

Obviously an inclusion $i : A \hookrightarrow X$ satisfies the **HEP** if every pair of maps $X \times \{0\} \rightarrow Y$ and $A \times I \rightarrow Y$ that agree on $A \times \{0\}$ can be extended to a map $X \times I \rightarrow Y$.

Lemma 5.4.3. A pair (X, A) is a cofibration iff $(X \times \{0\}) \cup (A \times I)$ is a retract of $X \times I$.

Proof. For one direction, the **HEP** for (X, A) implies that the identity map on $(X \times \{0\}) \cup (A \times I)$ extends to a map $X \times I \rightarrow (X \times \{0\}) \cup (A \times I)$, which makes $(X \times \{0\}) \cup (A \times I)$ a retract of $X \times I$.

The converse is equally easy when A is closed in X . Indeed, then any two maps $X \times \{0\} \rightarrow Y$ and $A \times I \rightarrow Y$ that agree on $A \times \{0\}$ combine to give a map $X \times \{0\} \cup A \times I \rightarrow Y$ which is

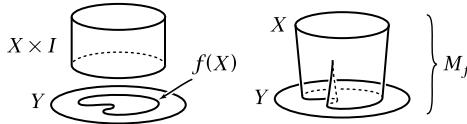


Fig. 5.1: Mapping cylinder of f .

continuous since it is continuous on those two closed sets separately. By composing this map with a retraction $X \times I \rightarrow X \times \{0\} \cup A \times I$ we get an extension $X \times I \rightarrow Y$.

The hypothesis that A is closed can be avoided by a more complicated argument, see [22, Prop. A.18]. For Hausdorff spaces X this is not needed, since if $(X \times \{0\}) \cup (A \times I)$ is a retract of $X \times I$ and X is Hausdorff, then A must in fact be closed in X . Indeed, if $r : X \times I \rightarrow X \times I$ is a retraction onto $(X \times \{0\}) \cup (A \times I)$, then the image of r is the set of points $z \in X \times I$ with $r(z) = z$, a closed set if X is Hausdorff, so $(X \times \{0\}) \cup (A \times I)$ is closed in $X \times I$ and hence A is closed in X . ■

The case of Hausdorff X will suffice for our needs, so the above proof is complete as far as we are concerned.

Example 5.4.4. A simple example of a pair (X, A) with A closed for which the HEP fails is the pair (I, A) where $A = \{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. There is no continuous retraction $I \times I \rightarrow (I \times \{0\}) \cup (A \times I)$. The breakdown of homotopy extension here can be attributed to the bad structure of the pair near 0.

Definition 5.4.5 (Mapping cylinder). Let $f : X \rightarrow Y$ be a map. The mapping cylinder of f is the pushout Z_f , see Figure 5.1, defined as the quotient of $(X \times I) \sqcup Y$ that identifies $(x, 0) \in X \times I$ with $f(x) \in Y$:

$$\begin{array}{ccc}
 X \sqcup X & \xrightarrow{f \sqcup \text{id}} & Y \sqcup X \\
 i_0 \sqcup i_1 \downarrow & & \downarrow s \sqcup j, \\
 X \times I & \xrightarrow{a} & Z_f
 \end{array}
 \quad Z_f = (X \times I) \sqcup Y / (x, 0) \sim f(x), \quad (5.4.2)$$

$$s(y) = y, \quad j(x) = (x, 1), \\
 i_0(x) = (x, 0), \quad i_1(x) = (x, 1).$$

We also have the projection $q : Z_f \rightarrow Y$, $(x, t) \mapsto f(x)$, $y \mapsto y$.

The images $q(X \times \{1\})$ and $q(Y)$ are homeomorphic to X and Y respectively. When X and Y are homotopy equivalent, both of these subspaces of Z_f can be shown to be deformation retracts of Z_f . Therefore two spaces X, Y are homotopy equivalent iff they are homeomorphic

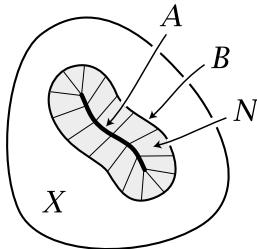


Fig. 5.2: Mapping cylinder neighborhood from Example 5.4.6.

to deformation retracts of some third space Z . See e.g. [8, Prop. 7.46] for a proof. We will also provide a proof in Corollary 5.4.12.

Example 5.4.6. A pair (X, A) has the **HEP** if A has a mapping cylinder neighborhood of X , by which we mean a closed neighborhood N containing a subspace B , thought of as the boundary of N , with $N \setminus B$ an open neighborhood of A , such that there exists a map $f : B \rightarrow A$ and a homeomorphism $h : Z_f \rightarrow N$ with $h|_{A \cup B} = \text{id}$. Mapping cylinder neighborhoods like this occur fairly often. For example thickened graphs in the plane, see Figure 5.2.

Exercise 5.2. Show that the mapping cylinder of the double covering map π of the circle \mathbb{S}^1 ($\pi(z) = z^2$ on the unit circle in \mathbb{C}) is homeomorphic to a bounded Möbius band.

Proposition 5.4.7 ([22, Prop. 0.16]). *If (X, A) is a CW pair (i.e. A is a subcomplex of X), then $(X \times \{0\}) \cup (A \times I)$ is a deformation retract of $X \times I$, hence (X, A) is a cofibration.*

Proof. There is a retraction $r : \mathbb{D}^n \times I \rightarrow (\mathbb{D}^n \times \{0\}) \cup (\partial \mathbb{D}^n \times I)$, for example the radial projection from the point $(0, 2) \in \mathbb{D}^n \times \mathbb{R}$. Then setting $r_t = t \cdot r + (1-t) \text{id}$ gives a deformation retraction of $\mathbb{D}^n \times I$ onto $(\mathbb{D}^n \times \{0\}) \cup (\partial \mathbb{D}^n \times I)$. This deformation gives rise to a deformation retraction of $X^{(n)} \times I$ onto $(X^{(n)} \times \{0\}) \cup ((X^{(n-1)} \cup A^{(n)}) \times I)$ since $X^{(n)}$ is obtained from the latter by attaching copies of $\mathbb{D}^n \times I$ along $(\mathbb{D}^n \times \{0\}) \cup (\partial \mathbb{D}^n \times I)$. If we perform this deformation retraction in dimension n during the t -interval $[2^{-n-1}, 2^{-n}]$, this infinite concatenation of homotopies is a deformation retraction of $X \times I$ onto $(X \times \{0\}) \cup (A \times I)$. There is no problem with continuity of this deformation at $t = 0$ since it is continuous on $X^{(n)} \times I$, being stationary there during the t -interval $[0, 2^{-n-1}]$, and CW-complexes have the weak topology with respect to their skeleta so a map is continuous iff its restriction to each skeleton is continuous. ■

Proposition 5.4.8 ([22, Prop. 0.17]). *If the pair (X, A) is a cofibration and A is contractible, then the quotient map $q : X \rightarrow X/A$ is a homotopy equivalence.*

Proof. We can use the HEP to extend a contraction of A to a homotopy of maps $X \rightarrow X$ that map A to A . This descends to a homotopy of maps $X/A \rightarrow X/A$. At $t = 1$ this homotopy produces a map that maps A to the point to which A was contracted, so this map induces a map $X/A \rightarrow X$. It is then easy to see that this map is a homotopy inverse of q via the two homotopies constructed above. ■

Proposition 5.4.9 ([22, Prop. 0.18]). *If (X_1, A) is a CW-pair and we have the attaching maps $f, g : A \rightarrow X_0$ that are homotopic, then $X_0 \sqcup_f X_1 \simeq X_0 \sqcup_g X_1$ rel X_0 .*

Proof. If $F : A \times I \rightarrow X_0$ is the homotopy from f to g , then consider the space $X_0 \sqcup_F (X_1 \times I)$. This contains both spaces in the claim of the theorem as subspaces. A deformation retraction of $X_1 \times I$ onto $(X_1 \times \{0\}) \cup (A \times I)$ as in Proposition 5.4.7 induces a deformation retraction of $X_0 \sqcup_F (X_1 \times I)$ onto $X_0 \sqcup_f X_1$. Similarly that space retracts onto $X_0 \sqcup_g X_1$. Both of these deformation retractions restrict to the identity map on X_0 , so they give the necessary homotopy equivalence. ■

Proposition 5.4.10 ([22, Prop. 0.19]). *Suppose (X, A) and (Y, A) are cofibrations and $f : X \rightarrow Y$ is a homotopy equivalence with $f|_A = \text{id}_A$. Then f is a homotopy equivalence rel A .*

Proof. We only outline the proof, see [22, Prop. 0.19] for the full details. Let $g : Y \rightarrow X$ be a homotopy inverse of f . The idea is to construct a homotopy from g to a map g_1 such that $g_1|_A = \text{id}_A$. This is possible by taking the existing homotopy from $g \circ f$ to id_X , restricting it to a homotopy from $g|_A$ to id_A , and extending it to a homotopy of maps $Y \rightarrow X$. At $t = 1$ this homotopy will give the necessary g_1 . Then we check that $g_1 \circ f \simeq \text{id}_X$ rel A and that $f \circ g_1 \simeq \text{id}_Y$ rel A . ■

By applying the above to the inclusion map we immediately get the next Corollary.

Corollary 5.4.11 ([22, Cor. 0.20]). *If (X, A) is a cofibration and the inclusion $i : A \hookrightarrow X$ is a homotopy equivalence, then A is a deformation retract of X .*

Corollary 5.4.12 ([22, Cor. 0.21]). *A map $f : X \rightarrow Y$ is a homotopy equivalence iff X is a deformation retract of the mapping cylinder Z_f . Hence, two spaces $X \simeq Y$ iff there is a third space containing both X and Y as deformation retracts.*

Proof. We have the inclusions $i : X \hookrightarrow Z_f$ and $j : Y \hookrightarrow Z_f$, and the canonical retraction $r : Z_f \rightarrow Y$, so $f = r \circ i$ and $i \simeq j \circ f$. Since j and r are homotopy equivalences (“h.e.”), it follows that f is a h.e. iff i is h.e., since the composition of two h.e.’s is a h.e. and a map homotopic to a h.e. is also a h.e. Now we apply the preceding Corollary to the pair (Z_f, X) , which is a cofibration by Example 5.4.6 using the neighborhood $X \times [0, \frac{1}{2}]$ of X in Z_f . ■

5.5 Whitehead’s theorem

Let us first adapt the Compression Criterion (Proposition 5.2.3) to the context of CW-complexes.

Lemma 5.5.1 (Compression Lemma). *Let (X, A) be a CW pair (i.e. A is a subcomplex of X) and let (Y, B) be any pair with $B \neq \emptyset$. For each n such that $X \setminus A$ has cells of dimension n , assume that $\pi_n(Y, B, y_0) = 0$ for all $y_0 \in B$. Then every map $f : (X, A) \rightarrow (Y, B)$ is homotopic rel A to a map $X \rightarrow B$.*

When $n = 0$ the condition $\pi_n(Y, B, y_0) = 0$ for all $y_0 \in B$ is to be regarded as saying that (Y, B) is connected.

Proof. Assume inductively that f has already been homotoped to take the skeleton $X^{(k-1)}$ to B . If Φ is the characteristic map of a k -cell e^k of $X \setminus A$, the composition $f \circ \Phi : (\mathbb{D}^k, \partial \mathbb{D}^k) \rightarrow (Y, B)$ can be homotoped into B rel $\partial \mathbb{D}^k$ in view of the hypothesis that $\pi_k(Y, B, y_0) = 0$ if $k > 0$, or that (Y, B) is 0-connected if $k = 0$. This homotopy of $f \circ \Phi$ induces a homotopy of f on the quotient space $X^{(k-1)} \cup e^k$ of $X^{(k-1)} \sqcup \mathbb{D}^k$, a homotopy rel $X^{(k-1)}$. Doing this for all k -cells of $X \setminus A$ simultaneously, and taking a constant homotopy on A , we obtain a homotopy of $f|_{X^{(k)} \cup A}$ to a map into B . By the homotopy extension property, this homotopy extends to a homotopy defined on all of X , and the induction step is completed.

Finitely many applications of the induction step finish the proof if the cells of $X \setminus A$ are bounded in dimension. Otherwise, we can perform the homotopy of the induction step during the t -interval $[1 - 2^{-k}, 1 - 2^{-k-1}]$. Any finite skeleton $X^{(k)}$ is eventually stationary under these homotopies, hence we have a well-defined homotopy f_t , $t \in [0, 1]$ with $f_1(X) \subset B$. ■

Theorem 5.5.2 (Whitehead’s Theorem). *If a map $f : X \rightarrow Y$ between connected CW-complexes induces isomorphisms $f_* : \pi_n(X) \rightarrow \pi_n(Y)$ for all n , then f is a homotopy equivalence.*

If f is the inclusion map of a subcomplex $X \hookrightarrow Y$, the conclusion is stronger: X is a deformation retract of Y .

Proof. In the special case that f is the inclusion of a subcomplex, consider the long exact sequence on all homotopy groups of (Y, X) . Since f induces isomorphisms of all homotopy

groups, the relative groups $\pi_n(Y, X)$ are zero. Applying the Compression Lemma to the identity map on (Y, X) then yields a deformation retraction of Y onto X .

The general case can be proved using mapping cylinders. Recall that the mapping cylinder Z_f is the quotient space of the disjoint union of $X \times I$ and Y under the identifications $(x, 1) \sim f(x)$. Thus Z_f contains both $X = X \times 0$ and Y as subspaces, and Z_f deformation retracts onto Y . The map f becomes the composition of the inclusion $X \hookrightarrow Z_f$ with the retraction $Z_f \rightarrow Y$. Since this retraction is a homotopy equivalence, it suffices to show that Z_f deformation retracts onto X if f induces isomorphisms on homotopy groups, or equivalently, if the relative groups $\pi_n(Z_f, X)$ are all trivial.

If the map f happens to be *cellular*, i.e. taking $X^{(n)}$ to $Y^{(n)}$, then (Z_f, X) is a CW-pair and so we are done by the first paragraph of the proof. If f is not cellular, we can either appeal to the Cellular Approximation Theorem 5.6.2 which says that f is homotopic to a cellular map, or use the following argument. First apply the Compression Lemma to obtain a homotopy rel X of the inclusion $(X \cup Y, X) \hookrightarrow (Z_f, X)$ to a map into X . Since the pair $(Z_f, X \cup Y)$ obviously is a cofibration, this homotopy extends to a homotopy from id_{Z_f} to a map $g : Z_f \rightarrow Z_f$ taking $X \cup Y$ into X . Then apply the Compression Lemma again to the composition $((X \times I) \sqcup Y, (X \times \partial I) \sqcup Y) \xrightarrow{g} (Z_f, X \cup Y) \xrightarrow{g} (Z_f, X)$ to finish the construction of a deformation retraction of Z_f onto X . ■

Remark 5.5.3. The Whitehead theorem *does not* say that two complexes with isomorphic homotopy groups are homotopy equivalent – the existence of a map f that generates those isomorphisms is crucial! For example, $\mathbb{R}P^2$ and $\mathbb{S}^2 \times \mathbb{R}P^\infty$ both have fundamental group \mathbb{Z}_2 and trivial higher groups (since their universal covers \mathbb{S}^2 and $\mathbb{S}^2 \times \mathbb{S}^\infty$ are homotopy equivalent, \mathbb{S}^∞ being contractible). But these two spaces are not homotopy equivalent because they have vastly different homology groups.

Definition 5.5.4 (Weak homotopy equivalence). A map $f : X \rightarrow Y$ is called a weak homotopy equivalence if all the induced homomorphisms $f_* : \pi_n(X, *) \rightarrow \pi_n(Y, *), n \geq 0$, are isomorphisms regardless of the choice of basepoints.

Whitehead's theorem therefore states that a weak homotopy equivalence between CW-complexes is actually a homotopy equivalence. In general, however, weak homotopy equivalences are a much bigger class of maps. For example, there exist noncontractible spaces whose homotopy groups are all trivial, such as the ‘quasi-circle’. For such spaces a map to a point is a weak homotopy equivalence that but not a homotopy equivalence.

One special case where isomorphisms by themselves do imply homotopy equivalence is when all homotopy groups are trivial.

Definition 5.5.5 (Weakly contractible space). A space X is called weakly contractible if $\pi_n(X)$ is trivial for all $n \geq 0$.

Corollary 5.5.6. *A weakly contractible CW-complex is contractible.*

Proof. The inclusion map of a 0-cell into the complex induces an isomorphism of the trivial homotopy groups, therefore by Whitehead's theorem there is a homotopy equivalence with a point. ■

Lemma 5.5.7 (Extension Lemma). *Let (X, A) be a CW-pair and Y is a path-connected space such that $\pi_{n-1}(Y) = 0$ for all n in which $X^{(n)} \setminus A^{(n)} \neq \emptyset$. Then any continuous map $f : A \rightarrow Y$ can be extended to a map $\tilde{f} : X \rightarrow Y$.*

Proof. Assume f was already extended to a map $f^{(n-1)}$ over the $(n-1)$ -skeleton. Then an extension over all n -cells exists iff the composition of the cell's attaching map $\mathbb{S}^{n-1} \rightarrow X^{(n-1)}$ with $f^{(n-1)}$ is nullhomotopic. ■

5.6 Cellular approximation

Recall when we applied the van Kampen theorem to compute $\pi_1(\mathbb{S}^n) = 0$ for $n \geq 2$, we implicitly used the fact that every loop can be deformed to miss at least one point in \mathbb{S}^n , and then it can obviously be contracted to a point since $\mathbb{S}^n \setminus *$ is contractible. This is not an obvious fact since there exist space-filling loops that are surjective. Similar space-filling maps exist for all pairs $\mathbb{S}^m \rightarrow \mathbb{S}^n$ with $m < n$.

For homotopy theory of CW complexes it turns out to be sufficient to work with maps that map n -skeleta to n -skeleta, and any other map is homotopic to one of this class.

Lemma 5.6.1 ([22, Lem. 4.10]). *Let $f : Y^n \rightarrow Z$ be a map, where Z is obtained from a subspace W by attaching a k -cell e^k . Then there is a homotopy $f_t : (I^n, f^{-1}(e^k)) \rightarrow (Z, e^k)$ rel $f^{-1}(W)$ from $f = f_0$ to a map f_1 for which there is a polyhedron $K \subset I^n$ such that:*

- (a) *$f_1(K) \subset e^k$ and $f_1|_K$ is piecewise linear (i.e. linear on each face of K) with respect to some identification of e^k with \mathbb{R}^n .*
- (b) *There exists a nonempty open $U \subset e^k$ such that $f_1^{-1}(U) \subset K$.*

Proof. Identifying e^k with \mathbb{R}^k , let $B_1, B_2 \subset e^k$ be the closed balls of radius 1 and 2 at the origin. Since $f^{-1}(B_2)$ is closed and therefore compact in I^n , it follows that f is uniformly continuous on $f^{-1}(B_2)$. Thus there exists $\epsilon > 0$ such that $|x - y| < \epsilon$ implies $|f(x) - f(y)| < \frac{1}{2}$ for all $x, y \in f^{-1}(B_2)$. Subdivide the interval I so that the induced subdivision of I^n into cubes has each cube lying in a ball of diameter less than ϵ . Let K_1 be

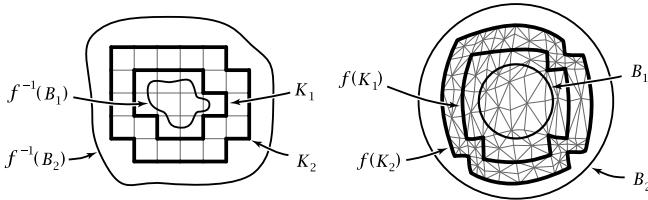


Fig. 5.3: Subdivision from the proof of Lemma 5.6.1.

the union of all the cubes meeting $f^{-1}(B_1)$ and let K_2 be the union of all cubes meeting K_1 . We may assume ϵ is chosen smaller than half the distance between the compact sets $f^{-1}(B_1)$ and $I^n \setminus f^{-1}(\bigcap B_2)$, and then we will have $K_2 \subset f^{-1}(B_2)$. See Figure 5.3.

Now we subdivide all the cubes of K_2 into simplices. This can be done inductively in the dimension by taking the center point of a cube as a new vertex and joining it by a cone to each simplex in boundary of the cube.

Let $g : K_2 \rightarrow e^k = \mathbb{R}^k$ be the map that equals f on all vertices of simplices of the subdivision and is linear on each simplex. Let $\varphi : K_2 \rightarrow [0, 1]$ be the map that is linear on simplices and has value 1 on vertices in K_1 and 0 on vertices of $K_2 \setminus K_1$. Thus $\varphi|_{K_1} = 1$. Define a homotopy $f_t : K_2 \rightarrow e^k$ by the formula $(1 - t\varphi) \cdot f + (t\varphi) \cdot g$, so $f_0 = f$ and $f_1|_{K_1} = g|_{K_1}$. Since f_t is constant in t on simplices in K_2 disjoint from K_1 , and in particular on simplices in the closure $\overline{I^n \setminus K_2}$, we may extend f_t to be the constant homotopy of f on $I^n \setminus K_2$.

The map f_1 takes the closure $\overline{I^n \setminus K_1}$ to a compact set C which, we claim, is disjoint from the center 0 of B_1 and hence from a neighborhood U of 0. This will prove the lemma, with $K = K_1$ since we will have $f_1^{-1}(U) \subset K_1$ with f_1 piecewise linear on K_1 where it is equal to g .

The verification of the claim has two steps:

- (1) On $I^n \setminus K_2$ we have $f_1 = f$, and f takes $I^n \setminus K_2$ outside B_1 since $f^{-1}(B) \subset K_2$ by construction.
- (2) For a simplex σ of K_2 not in K_1 we have $f(\sigma)$ contained in some ball B_σ of radius $\frac{1}{2}$ by the choice of ϵ and the fact that $K_2 \subset f^{-1}(B_2)$. Since B_σ is convex, we must also have $g(\sigma) \subset B_\sigma$. We know that $g(\sigma) \subset B_\sigma$, hence also $f_1(\sigma) \subset B_\sigma$ for all t , and in particular $f_1(\sigma) \subset B_\sigma$. We know that $B_\sigma \not\subset B_1$ since $\sigma \not\subset K_1$ and thus $\sigma \not\subset f^{-1}(B_1)$. The radius of B_σ is half that of B_1 , so it follows that $0 \notin B_\sigma$, and hence $0 \notin f_1(\sigma)$. ■

Theorem 5.6.2 (Cellular Approximation Theorem). *Every map $f : X \rightarrow Y$ of CW-complexes is homotopic to a cellular map.*

Every map $f : (X, A) \rightarrow (Y, B)$ of CW-pairs is homotopic rel A to a cellular map.

Proof. Suppose inductively that f is cellular on $X^{(n-1)}$, and let e^n be an n -cell of X . The closure of e^n in X is compact, being the image of a characteristic map for e^n (Proposition 2.1.40), so f takes the closure of e^n to a compact set in Y . Therefore $f(e^n)$ meets finitely many cells of Y . Let $e^k \subset Y$ be a cell of highest dimension meeting $f(e^n)$. We may assume $k > n$, otherwise f is already cellular on e^n .

Now we apply Lemma 5.6.1 to the composition of $f : X^{(n-1)} \cup e^n \rightarrow Y^{(k)}$ with a characteristic map $I^n \rightarrow X$ for e^n , with $Z = Y^{(k)}$ and $W = Y^{(k)} \setminus e^k$. The homotopy given by the lemma is fixed on ∂I^n , hence induces a homotopy f_t of $f|_{X^{(n-1)} \cup e^n}$ fixed on $X^{(n-1)}$. The image of the resulting map f_1 intersects the open set $U \subset e^k$ in a set contained in the union of finitely many hyperplanes of dimension at most n , so if $n < k$ there will be points $p \in U$ not in the image of f_1 .

In summary, it is possible to deform $f|_{X^{(n-1)} \cup e^n}$, staying fixed on $X^{(n-1)}$, so that $f(e^n)$ misses some point $p \in e^k$. Then we can deform $f|_{X^{(n-1)} \cup e^n}$ rel $X^{(n-1)}$ so that $f(e^n)$ misses the whole cell e^k by composing with a deformation retraction of $Y^{(k)} \setminus \{p\}$ onto $Y^{(k)} \setminus e^k$. By finitely many iterations of this process we eventually make $f(e^k)$ miss all cells of dimension greater than n . Doing this for all n -cells simultaneously, staying fixed on n -cells in A where f is already cellular, we obtain a homotopy of $f|_{X^{(n)}}$ rel $X^{(n-1)} \cup A^{(n)}$ to a cellular map.

The induction step is then completed by appealing to the HEP in Proposition 5.4.7 to extend this homotopy, together with the constant homotopy on A , to a homotopy defined on all of X . Letting n go to infinity, the resulting possibly infinite sequence of homotopies can be realized as a single homotopy by performing the n -th homotopy during the t -interval $[1 - 2^{-n}, 1 - 2^{-n-1}]$. This homotopy will eventually be stationary on every k -skeleton and hence continuous at 1. ■

Corollary 5.6.3. $\pi_n(\mathbb{S}^k) = 0$ for $n < k$.

Proof. If \mathbb{S}^n and \mathbb{S}^k are given CW structures with one 0-cell (serving as basepoint) and one $n(k)$ -cell, then every pointed map $\mathbb{S}^n \rightarrow \mathbb{S}^k$ can be homotoped, fixing the basepoint, to be cellular, and hence constant if $n < k$. ■

5.7 CW approximation

Whitehead's theorem obviously holds not just for CW-complexes, but also for spaces that are homotopy equivalent to them. We now show that any topological space is *weakly* homotopy

equivalent to a CW-complex.

Definition 5.7.1 (CW approximation). A CW approximation of a topological space X is a CW-complex Z with a weak homotopy equivalence $f : Z \rightarrow X$.

Proposition 5.7.2 ([22, Prop. 4.13]). *Every space X has a CW approximation $f : Z \rightarrow X$. If X is path-connected, Z can be chosen to have a single 0-cell, with all other cells attached by pointed maps. Thus every connected CW-complex is homotopy equivalent to a CW-complex with these additional properties.*

Proof. The construction of a CW approximation $f : Z \rightarrow X$ for a space X is inductive. Suppose given a CW-complex A with a map $f : A \rightarrow X$ and suppose we have chosen a basepoint 0-cell $*$ in each component of A . Then for an integer $k \geq 0$ we will attach k -cells to A to form a CW complex B with a map $f : B \rightarrow X$ extending the given f , such that the induced homomorphism $f_* : \pi_i(B, *) \rightarrow \pi_i(X, f(*))$ is injective for $i = k - 1$ and surjective for $i = k$, for all basepoints $*$.

There are two steps:

- (1) Choose maps $\varphi_\alpha : (\mathbb{S}^{k-1}, s_0) \rightarrow (A, *)$ representing a set of generators for the kernel of $f_* : \pi_{k-1}(A, *) \rightarrow \pi_{k-1}(X, f(*))$ for all basepoints $*$. We may assume the maps φ_α are cellular, where \mathbb{S}^{k-1} has its standard CW structure with s_0 as the 0-cell. Attaching cells e_α^k to A via φ_α then produces a CW complex, and the map f extends over these cells using nullhomotopies of $f \circ \varphi_\alpha$, which exist by the choice of the φ_α 's.
- (2) Choose maps $f_\beta : \mathbb{S}^k \rightarrow X$ representing generators for the groups $\pi_k(X, f(*))$, attach cells e_β^k to A via the constant maps at the appropriate basepoints $*$, and extend f over the resulting spheres S_β^k via the f_β 's.

The surjectivity of f_* holds by construction. For the injectivity, an element of the kernel of $f_* : \pi_{k-1}(B, *) \rightarrow \pi_{k-1}(X, f(*))$ can be represented by a cellular map $h : \mathbb{S}^{k-1} \rightarrow B$. This has image in A , so is in the kernel of $f_* : \pi_{k-1}(A, *) \rightarrow \pi_{k-1}(X, f(*))$ and hence is homotopic to a linear combination of the φ_α 's, which are nullhomotopic in B , so h is nullhomotopic as well. When $k = 1$ there is no group structure on π_{k-1} so injectivity on π_0 does not follow from having a trivial kernel, and we modify the construction by choosing the cells e_α^1 to join each pair of basepoints $*$ that map by f to the same path-component of X . The map f can then be extended over these 1-cells e_α^1 .

Note that if f happened to be injective or surjective on π_i for some $i < k - 1$ or $i < k$, respectively, then this remains true after attaching the k -cells. This is because attaching k -cells does not affect π_i if $i < k - 1$, by cellular approximation, nor does it destroy surjectivity on π_{k-1} or indeed any π_i , obviously.

Now to construct a CW approximation $f : Z \rightarrow X$ one can start with A consisting of one point for each path-component of X , with $f : A \rightarrow X$ mapping each of them to the corresponding

path-component. Having now a bijection on π_0 , attach 1-cells to A to create a surjection on π_1 for each path-component, then 2-cells to improve this to an isomorphism on π_1 and a surjection on π_2 , and so on for each successive π_i . After all cells have been attached one has a CW-complex Z and a weak homotopy equivalence $f : Z \rightarrow X$. ■

Remark 5.7.3. One can apply this technique to produce CW approximations of pairs (X, A) . First construct a CW approximation $f_A : Z_A \rightarrow A$, then starting with the composition $Z_A \rightarrow A \hookrightarrow X$, attach cells to Z_A to create a weak homotopy equivalence $f : Z \rightarrow X$ extending f_A . The induced map f_* will be an isomorphism of relative homotopy groups by the five-lemma 22.3.3 applied to the two sequences of morphisms in relative and absolute homotopy.

Running ahead, we will be calling a CW-pair (Z, A) *n-connected* if $\pi_k(Z, A, a) = 0$ for $k = 1, \dots, n$ and all $a \in A$, and the inclusion-induced map $\pi_0(A) \rightarrow \pi_0(Z)$ is surjective.

Definition 5.7.4 (*n*-connected CW model). An *n-connected CW model* for a CW pair (X, A) with a nonempty subcomplex A is an *n-connected CW pair* (Z, A) and a map $f : Z \rightarrow X$ with $f|_A = \text{id}_A$ such that the induced homomorphism $f_* : \pi_i(Z) \rightarrow \pi_i(X)$ is an isomorphism for $i > n$ and an injection for $i = n$, for all choices of basepoint.

One can think of Z as a homotopy-theoretic hybrid of X and A . As n increases it looks more and more like A , and less like X . We have just shown that *n-connected models* always exist, and can be obtained from A by attaching cells in dimensions $> n$ so that by cellular approximation (Z, A) is *n-connected*. Moreover, it is not difficult to show ([22, Prop. 4.18, Cor. 4.19]) that these models (and therefore CW approximations as well) are unique up to homotopy equivalence rel A . In particular, CW approximations are unique up to homotopy equivalence.

Example 5.7.5 (Postnikov towers). We can use this technique to construct, for each connected CW-complex X , and each $n \geq 1$, a CW-complex X_n containing X as a subcomplex such that:

- (a) $\pi_i(X_n) = 0$ for $i > n$;
- (b) The inclusion $X \hookrightarrow X_n$ induces an isomorphism on π_i for $i \leq n$.

To do this, apply the general construction to the constant map of X to a point, starting at the stage of attaching cells of dimension $n + 2$. Thus we attach $(n + 2)$ -cells to X using cellular maps $\mathbb{S}^{n+1} \rightarrow X$ that generate $\pi_{n+1}(X)$ to form a space with trivial π_{n+1} , then for this space we attach $(n + 3)$ -cells to make π_{n+2} trivial, and so on. The result is a complex X_n with the desired properties.

The inclusion $X \hookrightarrow X_n$ extends to a map $X_{n+1} \rightarrow X_n$ since X_{n+1} is obtained from X by attaching cells of dimension at least $n + 3$, and $\pi_i(X_n) = 0$ for $i > n$ so we can apply the Extension Lemma 3.3.4. Thus we have a commutative diagram called a Postnikov tower for X . One can regard the spaces X_n as truncations of X which provide successively better

approximations to X as n increases. There is a weak homotopy equivalence between X and the inverse limit of the Postnikov tower $\varprojlim X_n$, i.e. X is a CW approximation of $\varprojlim X_n$.

$$\begin{array}{c}
 & & & \vdots \\
 & & & \downarrow \\
 & & X_3 & \\
 & & \downarrow & \\
 & X_2 & & \\
 & \nearrow & \downarrow & \\
 X & \longrightarrow & X_1 &
 \end{array} \tag{5.7.1}$$

Now we prove one of the statements announced in Remark 4.4.19.

Proposition 5.7.6. *For every (discrete) group G there exists a path-connected two-dimensional CW-complex X_G such that $\pi_1(X_G) \cong G$.*

Proof. Consider a presentation of the group as the quotient of a free group $\langle J \rangle$ by some normal subgroup $\langle R \rangle$ generated by a set of relations R . Construct the 1-skeleton $X^{(1)}$ of X_G as a wedge sum of circles $\bigvee_{j \in J} \mathbb{S}^1$. Every relation describes a loop in $X^{(1)}$: for example if $ab^{-1}c^2$ is a relation, then the loop is given by going around a once, around b once in the opposite direction, and then twice around c . For each relation $r \in R$ take a 2-cell \mathbb{D}^2 with boundary \mathbb{S}^1 and define an attaching map $\varphi_r : \mathbb{S}^1 \rightarrow X^{(1)}$ so that it represents the loop specified by r . The resulting CW-complex $X_G = X_G^{(2)}$ is then a 2-dimensional complex whose fundamental group is, by construction, isomorphic to G . ■

We will eventually incorporate this space X_G as into the so called classifying space BG , so $\pi_1(BG) \cong \pi_1(X_G) \cong G$, i.e. the inclusion is a weak 1-equivalence. The classifying space BG must have $\pi_{n+1}(BG) \cong \pi_n(G)$, so in the case of discrete G this is $\pi_{\geq 2}(BG) \cong 0$. In other words, BG for discrete groups is the first Postnikov truncation of the space X_G we just constructed.

Example 5.7.7. Consider the cyclic group $G = C_n \cong \mathbb{Z}/n\mathbb{Z}$ for some integer n . Then X_G is $\mathbb{R}P^2$ for $n = 2$, and for higher n it is a subcomplex of the lens space with fundamental group C_n .

Definition 5.7.8 (Aspherical space). A space X is called aspherical if $\pi_n(X) = 0$ for all $n \geq 2$.

Definition 5.7.9 (First Eilenberg–MacLane space). For a group G , the first Eilenberg–MacLane space for G , denoted $K(G, 1)$, is defined up to weak homotopy equivalence as an aspherical space whose fundamental group is G . It can be realized as the first Postnikov truncation $(X_G)_1$ of the space X_G constructed in Proposition 5.7.6, and in the case of discrete G it coincides with the so called classifying space BG . Since π_1 is the only nontrivial homotopy group of $K(G, 1)$, its universal covering space is contractible by Whitehead’s theorem.

Remark 5.7.10. We will later be able to define Eilenberg–MacLane and classifying spaces so that they are unique up to homotopy equivalence, not just weak homotopy equivalence. Namely they will be required to be quotients of weakly contractible CW -complexes by a free group action, and of course weakly contractible CW -complexes are actually contractible by Whitehead’s theorem, which will induce a homotopy equivalence between the resulting quotients as well.

Corollary 5.7.11. *For every discrete group G there exists a (unique up to homotopy equivalence) universal G -principal covering $EG \rightarrow K(G, 1)$ with contractible total space EG .*

Example 5.7.12 (Lens spaces [22, Example 2.43]). Given an integer $m > 1$ and integers l_1, \dots, l_n relatively prime to m , define the lens space $L = L_m(l_1, \dots, l_n)$ as the orbit space $\mathbb{S}^{2n-1}/\mathbb{Z}_m$ of the unit sphere $\mathbb{S}^{2n-1} \subset \mathbb{C}^n$ with the action of \mathbb{Z}_m generated by the rotation $(z_1, \dots, z_n) \mapsto (e^{2\pi i l_1/m} z_1, \dots, e^{2\pi i l_n/m} z_n)$. In particular, when $m = 2$, for any odd l_i ’s this is the antipodal map and $L = L_2(1, \dots, 1) = \mathbb{R}P^{2n-1}$. In general, the projection $\mathbb{S}^{2n-1} \rightarrow L$ is a covering space since the action of \mathbb{Z}_m is free: each point of \mathbb{S}^{2n-1} has at least some coordinate z_j nonzero and then $e^{2\pi i l_j/m} z_j \neq z_j$ for $0 < k < m$ because of the assumption that l_j is relatively prime to m .

One can construct a CW -structure on L with one cell in each dimension from 0 to $2n - 1$, see [22, Example 2.43]. These structures also induce a CW -structure on the infinite lens spaces $L_m(l_1, l_2, \dots) = \mathbb{S}^\infty/\mathbb{Z}_m$ defined in the same way, which equals the union of the spaces $L_m(l_1, l_2, \dots, l_n)$ for $n = 1, 2, \dots$, each of which is a subcomplex of the next. Since the universal covering space \mathbb{S}^∞ is contractible, the infinite lens space is an Eilenberg–MacLane space $K(\mathbb{Z}_m, 1)$.

Exercise 5.3. Show that the above definition of $L_p(q)$ agrees with the definition of $L(p; q)$ from Example 2.5.32. See illustrations [here](#). We will give a more direct solution via bundles in Example 15.2.7.

5.8 Hurewicz fibrations

Definition 5.8.1 (Hurewicz fibration). A continuous map $p : E \rightarrow B$ is a (Hurewicz) fibration if it satisfies the **Homotopy Lifting Property (HLP)** for all spaces X . That is, in the diagram (5.8.1) a “lifted” homotopy \tilde{H} with initial condition $\tilde{H}(x, 0) = \tilde{H}_0(x)$ that makes the diagram commute exists for any space X , any homotopy $h : X \times I \rightarrow B$, and any continuous $\tilde{H}_0 : X \rightarrow E$. Here, i_0 is the inclusion map $i_0(x) = (x, 0)$.

$$\begin{array}{ccc} X & \xrightarrow{\tilde{H}_0} & E \\ i_0 \downarrow & \nearrow \tilde{H} & \downarrow p \\ X \times I & \xrightarrow{H} & B \end{array} \quad (5.8.1)$$

A morphism between two fibrations $p_1 : E_1 \rightarrow B_1$ and $p_2 : E_2 \rightarrow B_2$ is a pair of continuous maps $f : B_1 \rightarrow B_2$ and $F : E_1 \rightarrow E_2$ such that the square formed by them commutes, $p_2 \circ F = f \circ p_1$. Such maps are called *fiber-preserving*.

Example 5.8.2. The natural projections in a direct product are fibrations. Indeed, for the projection $\pi : B \times F \rightarrow B$ the homotopy lifting problem, defined by a map $f : X \times I \rightarrow B$ and an initial condition $\tilde{f}_0 : X \times \{0\} \rightarrow B \times F$, is solved by the map

$$\tilde{f} : X \times I \rightarrow B \times F, \quad \tilde{f}(x, t) = (f(x, t), \tilde{f}_0(x, 0)). \quad (5.8.2)$$

Lemma 5.8.3. *The evaluation map $\epsilon = \epsilon_0 \times \epsilon_1 : B^I \rightarrow B \times B$, $\gamma \mapsto (\gamma(0), \gamma(1))$ is a fibration for any space B .*

Proof. Indeed, any map $X \times I \rightarrow B \times B$ can be lifted into B^I because this is equivalent to finding an extension $X \times I \times I \rightarrow B$. Here, the “initial condition” fixes the values of this extension on three sides of the square $I \times I$, namely on $X \times J^1 = (X \times \{0\} \times I) \cup (X \times I \times \{0, 1\})$, where the values on the first piece are given by the initial condition of the homotopy lifting problem, and on the second piece by the given map $X \times I \rightarrow B \times B$. But this space, $X \times J^1$, is a retract of $X \times I^2$, so we can extend the map simply by precomposing it with the retraction $r : X \times I^2 \rightarrow X \times J^1$ to get the required extension. ■

Definition 5.8.4 (Mapping fibration). Let $f : Y \rightarrow B$ be any continuous map. Define the space P_f , also known as the *mapping path space*, as the fibered product (pullback)

$$P_f = B^I \times_B Y = \{(\gamma, y) \in B^I \times Y \mid \gamma(1) = f(y)\}. \quad (5.8.3)$$

It results from the pullback square

$$\begin{array}{ccc} P_f & \longrightarrow & Y \\ \downarrow & & \downarrow f \\ B^I & \xrightarrow{\epsilon_1} & B \end{array} \quad (5.8.4)$$

The mapping fibration of f is the map $p_f : P_f \rightarrow B$ given by $p_f(\gamma, y) = \gamma(0)$.

Proposition 5.8.5. *For any map $f : Y \rightarrow B$, the mapping fibration $p_f : P_f \rightarrow B$ is indeed a fibration.*

Proof. Suppose we have a diagram (5.8.1) with $E = P_f$. Denote pr_1 and pr_2 the natural projections onto the factors of the direct product $B^I \times Y$. We can first extend $\text{pr}_2 \circ \tilde{H}_0$ as in:

$$\begin{array}{ccc} X & \xrightarrow{\text{pr}_2 \circ \tilde{H}_0} & Y \\ i_0 \downarrow & \nearrow \tilde{h}_0 & \\ X \times I & & \end{array} \quad (5.8.5)$$

and then use Lemma 5.8.3 to find a diagonal mapping as in

$$\begin{array}{ccc} X & \xrightarrow{\text{pr}_1 \circ \tilde{H}_0} & B^I \\ i_0 \downarrow & \nearrow \chi & \downarrow \epsilon \\ X \times I & \xrightarrow[H \times (f \circ \tilde{h}_0)]{} & B \times B \end{array} \quad (5.8.6)$$

We can verify that $\tilde{H} = \chi \times \tilde{h}_0 : X \times I \rightarrow P_f$ is the required lifting. Indeed it lands in P_f because $\epsilon_1 \circ \chi = f \circ \tilde{h}_0$, and makes the diagram commute because $p \circ \chi = \epsilon_0 \circ \chi = H$ and $(\chi \times \tilde{h}_0) \circ i_0 = (\text{pr}_1 \circ \tilde{H}_0) \times (\text{pr}_2 \circ \tilde{H}_0) = \tilde{H}_0$. ■

As a consequence, we observe that every continuous map is a fibration up to homotopy equivalence on the domain.

Corollary 5.8.6. *Let $f : Y \rightarrow B$ be a continuous map. Denote by $\kappa_b \in B^I$ the constant path*

$\gamma(t) \equiv b$ for $b \in B$. Then the map ϕ in

$$\begin{array}{ccc} & P_f & \\ \phi \nearrow & \nwarrow & \downarrow p_f \\ Y & \xrightarrow{f} & B \end{array} \quad (5.8.7)$$

given by $\phi(y) = (\kappa_{f(y)}, y)$ is a homotopy equivalence that makes the diagram commute.

Proof. The homotopy in question restricts paths $\gamma : I \rightarrow B$ to the interval $[t, 1]$, so $\gamma \mapsto \gamma|_{[t, 1]}$. ■

Motivated by this, one defines the *homotopy fiber* of f over $b \in B$:

$$i : p_f^{-1}(b) = \{(\gamma, y) \mid \gamma(0) = b, \gamma(1) = f(y)\}. \quad (5.8.8)$$

Note that there is a canonical injection from the fiber of f into the fiber of p_f given by the restriction of ϕ , namely:

$$f^{-1}(b) \rightarrow p_f^{-1}(b), \quad y \mapsto (\kappa_b, y). \quad (5.8.9)$$

Thus the fiber is naturally embedded into the homotopy fiber, which is a ‘relaxed’ version of the fiber: the condition imposed on $y \in Y$ to sit in the fiber over b is that it has to be mapped to b by f , whereas a point (γ, y) in the homotopy fiber is a point $y \in Y$ together with a path “witnessing” that y lies in the fiber $f^{-1}(b)$ “up to homotopy”.

Proposition 5.8.7 (Homotopy theorem for fibrations, [22, Prop. 4.61]). *For a fibration $p : E \rightarrow B$ the fibers $F_b = p^{-1}(b)$ over each path component of B are all homotopy equivalent.*

Proof. A path $\gamma : I \rightarrow B$ gives rise to a homotopy of maps $g_t : F_{\gamma(0)} \rightarrow B$ with $g_t(F_{\gamma(0)}) = \{\gamma(t)\}$. The inclusion $F_{\gamma(0)} \hookrightarrow E$ provides a lifting \tilde{g}_t , so by the HLP we have a homotopy $\tilde{g}_t : F_{\gamma(0)} \rightarrow E$ with $\tilde{g}_t(F_{\gamma(0)}) \subset F_{\gamma(t)}$ for all t . In particular, \tilde{g}_1 gives the map $L_\gamma : F_{\gamma(0)} \rightarrow F_{\gamma(1)}$. The association $\gamma \mapsto L_\gamma$ has the following properties which we will prove below:

- (a) If $\gamma \simeq \gamma'$ with fixed endpoints, then $L_\gamma \cong L_{\gamma'}$. In particular, the homotopy class of L_γ is independent of the choice of the lifting \tilde{g}_t of g_t ;
- (b) For a composition of paths $\gamma \bullet \gamma'$, $L_{\gamma \bullet \gamma'}$ is homotopic to the composition $L_{\gamma'} \circ L_\gamma$.

From these it follows that L_γ is a homotopy equivalence with homotopy inverse $L_{\bar{\gamma}}$, where $\bar{\gamma}$ is the inverse path of γ .

Note that a fibration has the **HLP** for pairs $(X \times I, X \times \partial I)$ since the pairs $(I \times I, I \times \{0\} \cup \partial I \times I)$ and $(I \times I, I \times \{0\})$ are homeomorphic, hence the same is true after taking products with X .

To prove (a), let $\gamma(s, t)$ be a homotopy from $\gamma(t)$ to $\gamma'(t)$. This determines a family $g_{st} : F_{\gamma(0)} \rightarrow B$ with $g_{st}(F_{\gamma(0)}) = \{\gamma(s, t)\}$. Let $\tilde{g}_{0,t}$ and $\tilde{g}_{1,t}$ be liftings defining L_γ and $L'_{\gamma'}$, and let $\tilde{g}_{s,0}$ be the inclusion $F_{\gamma(0)} \hookrightarrow E$ for all s . Using the **HLP** for the pair $F_{\gamma(0)} \times I, F_{\gamma(0)} \times \partial I$, we can extend these liftings to liftings $\tilde{g}_{s,t}$ for all $(s, t) \in I^2$. Restricting to $t = 1$ then gives a homotopy $L_\gamma \simeq L_{\gamma'}$.

Property (b) holds since for liftings \tilde{g}_t and \tilde{g}'_t defining L_γ and $L_{\gamma'}$ we obtain a lifting defining $L_{\gamma \bullet \gamma'}$ by taking \tilde{g}_{2t} for $0 \leq t \leq 1/2$ and \tilde{g}'_{2t-1} for $1/2 \leq t \leq 1$. ■

Definition 5.8.8 (Fiber homotopy equivalence). A morphism of two fibrations $f : E_1 \rightarrow E_2$ is called a fiber homotopy equivalence if there is a morphism $g : E_2 \rightarrow E_1$ such that both compositions $f \circ g$ and $g \circ f$ are homotopic to the identity through morphisms of fibrations.

Proposition 5.8.9 ([22, Prop. 4.65]). *For a fibration $p : E \rightarrow B$, the inclusion $i : E \hookrightarrow P_p$ is a fiber homotopy equivalence. In particular, each fiber $p^{-1}(b)$ is homotopy equivalent to the corresponding homotopy fiber $p_f^{-1}(b)$.*

Proof. This follows from the general fact that two fibrations that are homotopy equivalent via a morphism of fibrations are actually fiber homotopy equivalent. However, let us give a direct proof here.

We apply the **HLP** to the homotopy $g_t : P_p \rightarrow B$, $g_t(\gamma, e) = \gamma(t)$, with initial lifting $\tilde{g}(\gamma, e) = e$. The lifting $\tilde{g}_t : P_p \rightarrow E$ uses then the second coordinate of a homotopy $h_t : P_p \rightarrow P_p$ whose first coordinate is the restriction of paths to the interval $[t, 1]$. Since the endpoints of the paths γ are unchanged, h_t is fiber-preserving. We have $h_0 = \text{id}_{P_p}$, $h_1(P_p) \subset E$, and $h_t(E) \subset E$ for all t . If we let i denote the inclusion $E \hookrightarrow P_p$ as above, then $i \circ h_1 \simeq \text{id}_{P_p}$ via h_t , and $h_1 \circ i \simeq \text{id}_{P_p}$ via $h_t|_E$, so i is a fiber homotopy equivalence. ■

Definition 5.8.10 (Pullback fibration). Given a fibration $p : E \rightarrow B$ and a map $f : X \rightarrow B$, the pullback or induced fibration $f^*p : f^*E \rightarrow X$ is defined by setting $f^*E = \{(e, x) \in E \times X \mid f(x) = p(e)\}$ with the projections of f^*E onto E and X giving a commutative diagram

$$\begin{array}{ccc} f^*E & \longrightarrow & E \\ f^*p \downarrow & & \downarrow p \\ X & \xrightarrow{f} & B \end{array} \tag{5.8.10}$$

This is indeed a fibration because a homotopy $h : Y \times I \rightarrow X$ has the lifting $\tilde{H} = (\widetilde{f \circ H}) \times H$, where $\widetilde{f \circ H}$ is a lifting of $f \circ H$ which exists since p is a fibration.

Proposition 5.8.11 ([22, Prop. 4.62]). *Given a fibration $p : E \rightarrow B$ and a homotopy of maps $f_t : X \rightarrow B$, the pullback fibrations $f_0^*E \rightarrow X$ and $f_1^*E \rightarrow X$ are fiber homotopy equivalent.*

Proof. Let $F : X \times I \rightarrow B$ be the homotopy. The pullback fibration $F^*E \rightarrow X \times I$ contains f_0^*E and f_1^*E over $X \times \{0\}$ and $X \times \{1\}$. So it suffices to prove the following: for a fibration $p : E \rightarrow B \times I$, the restricted fibrations $E_s = p^{-1}(B \times \{s\}) \rightarrow B$ are all fiber homotopy equivalent for $s \in [0, 1]$.

To prove this assertion the idea is to imitate the construction of the homotopy equivalences L_γ in the proof of Proposition 5.8.7. A path $\gamma : [0, 1] \rightarrow I$ gives rise to a fiber-preserving map $L_\gamma : E_{\gamma(0)} \rightarrow E_{\gamma(1)}$ by lifting the homotopy $g_t : E_{\gamma(0)} \rightarrow B \times I$, $g_t(e) = (p(e), \gamma(t))$, starting with the inclusion $E_{\gamma(0)} \hookrightarrow E$. As before, one shows the two basic properties (a) and (b), nothing that in (a) the homotopy $L_\gamma \simeq L_{\gamma'}$ is fiber-preserving since it is obtained by lifting a homotopy $h_t : E_{\gamma(0)} \times [0, 1] \rightarrow B \times I$ of the form $h_t(e, u) = (p(e), -)$. From (a) and (b) it follows that L_γ is a fiber homotopy equivalence with inverse $L_{\bar{\gamma}}$. ■

Corollary 5.8.12. *A fibration $E \rightarrow B$ over a contractible base B is fiber homotopy equivalent to a product fibration $B \times F \rightarrow B$.*

Proof. The pullback of E by the identity map $B \rightarrow B$ is E itself, while the pullback by a constant map $B \rightarrow B$ is a product $B \times F$. ■

Corollary 5.8.13. *A fibration $E \rightarrow B$ over a locally contractible base B is locally homotopy equivalent to a trivial fibration $U \times F \rightarrow B$.*

Proof. This follows because the restriction of a fibration to any subset $A \subset B$ of the base space is itself a fibration and we can use the previous Corollary. ■

5.9 Serre fibrations

As we have seen in Theorem 4.1.7, covering maps are a special case of Hurewicz fibrations, and moreover liftings in them are unique. Hurewicz fibrations are a pretty restrictive class of spaces due to the strong lifting requirement. In practice, when we are mostly interested in computing invariants such as the homotopy groups, we only ever need to lift mappings from simple spaces such as disks or cubes. This motivates the definition of the more general class of Serre fibrations.

Definition 5.9.1 (Serre fibration). A continuous map $p : E \rightarrow B$ is a Serre fibration if it satisfies the **HLP** for all disks \mathbb{D}^n , $n \geq 0$ (or cubes I^n).

It is also useful to introduce the notion of the *relative HLP* with respect to a pair (X, A) , $A \subset X$. Given a lifting diagram as in the definition of **HLP** and a lifting \tilde{H}_A already defined on $A \times I$, the lifted homotopy \tilde{H} can be taken to agree with \tilde{H}_A on $A \times I$. In other words, we have the diagram

$$\begin{array}{ccc} X \times \{0\} \cup (A \times I) & \xrightarrow{\tilde{H}_0 \cup \tilde{H}_A} & E \\ i_0 \cup i \downarrow & \nearrow \tilde{H} & \downarrow p \\ X \times I & \xrightarrow{H} & B \end{array} \quad (5.9.1)$$

The following Proposition shows that instead of having to prepare spaces in the form $\mathbb{D}^n \times I$, we can always lift maps relatively from the pair $(\mathbb{D}^{n+1}, \mathbb{S}^n)$ directly into a Serre fibration.

Proposition 5.9.2 ([26, Prop. 3.2.4]). *A map $p : E \rightarrow B$ has the **HLP** for the disk \mathbb{D}^n iff it has the relative **HLP** w.r.t. the pair $(\mathbb{D}^{n+1}, \mathbb{S}^n)$, i.e. for each commutative diagram*

$$\begin{array}{ccc} \mathbb{D}^n \times \{0\} \cup \mathbb{S}^n \times I & \xrightarrow{h_0} & E \\ i \downarrow & \nearrow \tilde{H} & \downarrow p \\ \mathbb{D}^n \times I & \xrightarrow{H} & B \end{array} \quad (5.9.2)$$

there exists $\tilde{H} : \mathbb{D}^n \times I \rightarrow E$ that makes the diagram commute.

Proof. All we need is a reparametrization of \mathbb{D}^{n+1} that transforms the “initial disk” $\mathbb{D}^n \times \{0\}$ into the union of that disk and the sides of the cylinder $\mathbb{D}^n \times I$. Indeed, there exists such a homeomorphism $k : (\mathbb{D}^n \times I, \mathbb{D}^n \times \{0\} \cup \mathbb{S}^{n-1} \times I) \rightarrow (\mathbb{D}^n \times I, \mathbb{D}^n \times \{0\})$ of pairs, shown in Figure 5.4. ■

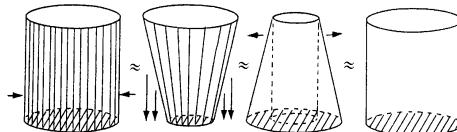


Fig. 5.4: Homeomorphism of pairs.

The next Theorem shows that a map $E \rightarrow B$ that is a Serre fibration locally around each point in B must be a Serre fibration. For Hurewicz fibrations the analogous statement is not generally true, i.e. a local Hurewicz fibration is not always a Hurewicz fibration, but holds, for example, over paracompact base spaces.

Theorem 5.9.3 ([26, Theorem 6.3.3]). *Let $p : E \rightarrow B$ be a continuous map and \mathcal{U} a collection of subsets whose interiors cover B . If for every $U \in \mathcal{U}$ the restriction $p_U : p^{-1}(U) \rightarrow U$ is a Serre fibration, then so is p .*

Proof. Subdivide the cube I^n into cubes of size δ and enumerate them I_i . We can choose δ so that each cube $I_i \times I$ is mapped inside a $U \in \mathcal{U}$ by the map H that we seek to lift. This is possible by the Lebesgue Lemma 4.1.4. Let $V^k \subset I^n$ denote the union of the k -dimensional faces of the subdivision of I^n .

We have to solve the lifting problem for the space I^n with initial condition h_0 . We begin by extending h_0 over $I^n \times [0, \delta]$ to a lifting of H . We solve the lifting problems

$$\begin{array}{ccc} I^n \times \{0\} \cup V^{k-1} \times [0, \delta] & \xrightarrow{\tilde{H}(k-1)} & E \\ i \downarrow & \nearrow \tilde{H}(k) & \downarrow p \\ I^n \times \{0\} \cup V^k \times [0, \delta] & \xrightarrow{H} & B \end{array} \quad (5.9.3)$$

for $k = 0, \dots, n$ with $V^{-1} = \emptyset$ and $\tilde{H}(-1) = a$ by induction in k . Let W be a k -dimensional cube and ∂W the union of its $(k-1)$ -dimensional faces. We can solve the lifting problems

$$\begin{array}{ccc} W \times \{0\} \cup \partial W \times [0, \delta] & \xrightarrow{\tilde{H}(k-1)} & p^{-1}(U) \\ i \downarrow & \nearrow \tilde{H}_W & \downarrow p_U \\ W \times [0, \delta] & \xrightarrow{H} & U \end{array} \quad (5.9.4)$$

by a map \tilde{H}_W , since p_U is a Serre fibration; here $U \in \mathcal{U}$ was chosen such that $H(W \times [0, \delta]) \subset U$.

The \tilde{H}_W 's for different $U \in \mathcal{U}$ combine to a continuous map $\tilde{H}(k) : V^l \times [0, \delta] \rightarrow E$ which covers H and extends $\tilde{H}(k-1)$. We define \tilde{H} on the first layer $I^n \times [0, \delta]$ as $\tilde{H}(n)$. We now treat $I^n \times [\delta, 2\delta]$ similarly with initial condition given by $\tilde{H}(n)|_{I^n \times \{\delta\}}$ and continue in this manner inductively. ■

Proposition 5.9.4 ([17, Prop. 3.2.3]). *Serre fibrations have the lifting property (without homotopy) with respect to all pairs of the form*

1. $(K \times I, (K \times \{0\}) \cup (L \times I))$, where K is a CW-complex and L is a subcomplex. This constitutes a relative **HLP** for the pair (K, L) .
2. (K, L) where K is a CW-complex and L is a subcomplex which is a strong deformation retract of K .

Proof. Let $\pi : E \rightarrow B$ be a Serre fibration, let K be a CW-complex and L a subcomplex of K .

1. Consider the lifting problem defined by some $f : K \times I \rightarrow B$ and an initial condition $\tilde{f}_0 : (K \times \{0\}) \cup (L \times I) \rightarrow E$. We prove the assertion by induction on the dimension k of the cells attached to L to build K . The case $k = 0$ is trivial. Thus, assume we have already constructed a lifting \tilde{f} over the subspace $(K^{(k)} \cup L) \times I \subset K \times I$ for some $k \geq 0$ and consider a $(k+1)$ -cell C not contained in L with characteristic map $\chi : \mathbb{D}^{k+1} \rightarrow K$. Since C is not contained in L , we have $C \cap (K^{(k)} \cup L) = C \cap L^{(k)}$. Hence, we wish to extend \tilde{f} from

$$(C \times \{0\}) \cup ((C \cap K^{(k)}) \times I) \subset C \times I \quad (5.9.5)$$

to a lifting of f on $C \times I$. Assume that we can extend

$$\tilde{f} \circ (\chi \times \text{id}_I) \Big|_{(\mathbb{D}^{k+1} \times \{0\}) \cup (\partial \mathbb{D}^{k+1} \times I)} \quad (5.9.6)$$

to a lifting of $f \circ (\chi \times \text{id}_I)$ on $\mathbb{D}^{k+1} \times I$. Since χ is injective on $\text{int } \mathbb{D}^{k+1}$, this lifting uniquely determines a lift of f on $C \times I$. By Proposition 5.3.4, applied to the CW-complex $C \times I$, the latter is continuous.

This argument shows that in order to prove that \tilde{f} extends to a lifting of f over $(K^{(k+1)} \cup L) \times I$, it suffices to show that π has the relative **HLP** with respect to the pair $(\mathbb{D}^{k+1}, \mathbb{S}^k)$. This holds by Proposition 5.9.2 since π is Serre fibration.

2. Let $F : K \times I \rightarrow K$ be a strong deformation retraction from K to L and consider the lifting problem defined by some $f : K \rightarrow B$ and an appropriate initial condition $\tilde{f}_0 : L \rightarrow E$. Define $g : K \times I \rightarrow B$ by $g = f \circ F$. Since F maps subsets $K \times \{1\}$ and $L \times I$ to L , we can also define

$$\tilde{g}_0 : (K \times \{1\}) \cup (L \times I) \rightarrow E, \quad \tilde{g}_0(x, t) = \tilde{f}_0(F(x, t)). \quad (5.9.7)$$

A brief calculation shows that \tilde{g}_0 is a lifting of g over the subset $(K \times \{1\}) \cup (L \times I)$. Hence, according to point 1, it can be extended to a lift \tilde{g} of g . Then, another brief calculation shows that the mapping $\tilde{f} : K \rightarrow E$ defined by $\tilde{f}(x) = \tilde{g}(x, 0)$ is a lifting of f through π extending \tilde{f}_0 . ■

Theorem 5.9.5. *Let $p : E \rightarrow B$ be a Serre fibration. For $B_0 \subset B$ let $E_0 = p^{-1}(B_0)$. Choose basepoints $* \in B_0$ and $* \in p^{-1}(*)$. Then p induces for $n \geq 1$ a bijection*

$$p_* : \pi_n(E, E_0, *) \rightarrow \pi_n(B, B_0, *).$$

Proof. First we show surjectivity. Let $g \in \pi_n(B, B_0, *)$ be represented by $\gamma : (\mathbb{D}^n, \mathbb{S}^{n-1}, *) \rightarrow (B, B_0, *)$. By HLP there exists a lifting $\tilde{\gamma} : \mathbb{D}^n \rightarrow E$ with $\tilde{\gamma}(*) = *$ and $p \circ \tilde{\gamma} = \gamma$. We then have $\tilde{\gamma}(\mathbb{S}^{n-1}) \subset E_0$ and thus $\tilde{\gamma}$ represents a pre-image pf g under p_* .

To show injectivity, let $g_0, g_1 \in \pi_n(E, E_0, *)$ be represented by γ_0, γ_1 and have the same image under p_* . Then there exists a homotopy $\phi_t : (\mathbb{D}^n, \mathbb{S}^{n-1}, *) \rightarrow (B, B_0, *)$ such that $\phi_0 = p \circ \gamma_0$ and $\phi_1 = p \circ \gamma_1$. Consider the subspace $T = \mathbb{D}^n \times \partial I \cup * \times I$ and define $G : T \rightarrow E$ by

$$G(x, t) = \begin{cases} \phi_t(x), & x \in \mathbb{D}^n, t \in \{0, 1\}, \\ *, & x = *, t \in I. \end{cases} \quad (5.9.8)$$

The set $T \subset \partial(\mathbb{D}^n \times I)$ is transformed into $*$ if one interchanges the last two coordinates. By HLP again, there exists a map $H : \mathbb{D}^n \times I \rightarrow E$ such that $H|_T = G$ and $p \circ H = \phi$. We can view H as a homotopy from γ_0 to γ_1 . ■

By taking B_0 to be a single point, we get the following corollary, which for $n \geq 2$ is just a generalization of Corollary 4.1.21 for covering spaces to all fibrations.

Corollary 5.9.6. *Given a Serre fibration $p : E \rightarrow B$ with $F = p^{-1}(*)$, there is an isomorphism of homotopy groups*

$$\pi_n(E, F, *) \cong \pi_n(B, *), \quad n \geq 1. \quad (5.9.9)$$

With the help of this isomorphism, we immediately get a special case of the long exact sequence of a pair (Theorem 5.2.9).

Theorem 5.9.7 (Exact sequence of a Serre fibration). *For a Serre fibration $p : E \rightarrow B$ with inclusion $i : F = p^{-1}(b) \subset E$ and $e \in F$ the sequence*

$$\cdots \rightarrow \pi_n(F, e) \xrightarrow{i_*} \pi_n(E, e) \xrightarrow{p_*} \pi_n(B, b) \xrightarrow{\partial} \pi_{n-1}(F, e) \rightarrow \cdots \rightarrow \pi_0(E, e) \rightarrow \pi_0(B, b) \quad (5.9.10)$$

is exact.

The boundary map ∂ has the following description. Let $f : (I^n, \partial I^n) \rightarrow (B, b)$ be given. View f as a map $I^{n-1} \times I \rightarrow B$ and lift it to $\tilde{f} : I^n \rightarrow E$, constant and equal to e on $J^{n-1} = I^{n-1} \times \{0\} \cup \partial I^{n-1} \times I$. Then $\partial[f]$ is represented by $\tilde{f}|_{I^{n-1} \times \{1\}}$. In plain language, a “loop” representing an element of $\pi_n(B)$ in the base can be lifted to a “path” in the total space whose “endpoints” lie in the same fiber F , and by restricting to F we get a lower-dimensional

“loop” that represents an element of $\pi_{n-1}(F)$. The very end of the sequence requires a little extra work.

Definition 5.9.8 (Trivialization). Let $\pi : E \rightarrow B$ be continuous and $U \subset B$ open and contained in the image of π . A *trivialization* of π over U is a homeomorphism $\varphi : \pi^{-1}(U) \rightarrow U \times F$ over U , i.e. a homeomorphism which satisfies $\text{pr}_1 \circ \varphi = \pi$ (where pr_1 is the natural projection onto the first component of the direct product). This condition determines F up to a homeomorphism since φ induces a homeomorphism of $\pi^{-1}(u)$ with $\{u\} \times F$.

Definition 5.9.9 (Locally trivial map). The map π is *locally trivial* if there exists an open covering \mathcal{U} of B such that π has a trivialization over each $U_\alpha \in \mathcal{U}$.

Note that a locally trivial map is open and hence a topological quotient map.

Example 5.9.10. Since a product projection is a fibration, we have by Theorem 5.9.3 that any locally trivial map is a Serre fibration.

Example 5.9.11 (Exact sequence of a covering space). Let $p : E \rightarrow B$ be a covering map with typical fiber F . Since each map $\mathbb{D}^n \rightarrow F$ must be constant because of discreteness of F , $\pi_n(F, *)$ is trivial for $n \geq 1$. The exact sequence of p then confirms the isomorphisms $\pi_n(E, f) \cong \pi_n(B, b)$, $n \geq 2$, that we proved in Corollary 4.1.21. Moreover we have the exact sequence (omitting the common basepoint)

$$1 \rightarrow \pi_1(E) \xrightarrow{p_*} \pi_1(B) \xrightarrow{\partial} \pi_0(F) \xrightarrow{i_*} \pi_0(E) \xrightarrow{p_*} \pi_0(B) \rightarrow 1 \quad (5.9.11)$$

with the inclusion $i : F = p^{-1}(*) \subset E$ and $\pi_0(F) = F$. It yields for the universal covering $p : \mathbb{R} \rightarrow \mathbb{S}^1$ the familiar bijection $\partial : \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$.

Definition 5.9.12 (Hopf fibrations). Looking slightly ahead, consider $\mathbb{S}^{2n-1} \subset \mathbb{C}^n$ as a space with a free action of $\mathbb{S}^1 = U(1)$ induced from scalar multiplication by $e^{2\pi it}$. Let U_j be the set of points $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ with $z_j \neq 0$. The map $z \mapsto z_j/|z_j|$ shows that $U_j \cong \mathbb{R}^{2n-1} \times \mathbb{S}^1$ is a trivial $U(1)$ -space. The orbit space of this action is exactly $\mathbb{C}P^{n-1} = \mathbb{S}^{2n-1}/\mathbb{S}^1$. The Hopf fibration is defined as the orbit map $p : \mathbb{S}^{2n-1} \rightarrow \mathbb{C}P^{n-1}$. It is indeed a fibration because it is clearly trivial over $U_j \cap \mathbb{S}^{2n-1}$. We will also see a more explicit definition of the most important $n = 2$ Hopf fibration later on.

Example 5.9.13. The exact sequence for the Hopf fibration, combined with $\pi_i(\mathbb{S}^1) = 0$ for $i > 1$ (see Corollary 4.1.22), yields the isomorphisms

$$p_* : \pi_i(\mathbb{S}^{2n+1}) \cong \pi_i(\mathbb{C}P^n), \quad i \geq 3; \quad (5.9.12)$$

and in particular $\pi_i(\mathbb{S}^3) \cong \pi_i(\mathbb{S}^2)$ for $i \geq 3$, since $\mathbb{C}P^1 \cong \mathbb{S}^2$.

Proposition 5.9.14 ([26, Prop. 6.3.8]). *Let $p : (E_1, E_0) \rightarrow B$ be a relative Serre fibration, i.e. $p : E_1 \rightarrow B$ is a Serre fibration and the restriction $p|_{E_0}$ is also a Serre fibration. Let (F_b^1, F_b^0) be the pair of fibers over $p_1(e) = b \in B$. Then:*

1. *The inclusion induces bijections $\pi_n(F_b^1, F_b^0, e) \cong \pi_n(E_1, E_0, e)$.*
2. *$\pi_0(E_0) \rightarrow \pi_0(E_1)$ is surjective iff $\pi_0(F_b^0) \rightarrow \pi_0(F_b^1)$ is surjective for each $b \in B$.*

Proof. (1) We first prove the claim for $n = 1$ and begin with surjectivity. Let $f : (I, \partial I, 0) \rightarrow (E_1, E_0, e)$ be given. The path $\overline{p \circ f} : I \rightarrow B$ (recall that the bar represents reversing the path) is lifted to $g : I \rightarrow E_0$ with initial point $f(1)$. Then $g(1) \in F_b^0$, and f and $f \bullet g$ represent the same element in $\pi_1(E_1, E_0, e)$. The projection $p \circ (f \bullet g)$ is a nullhomotopic loop with basepoint b . We lift a nullhomotopy to E_1 with initial condition $f \bullet g$ on $I \times \{0\}$ and constant on $\partial I \times I$. The lifting is a homotopy $(I, \partial I, 0) \times I \rightarrow (E_1, E_0, e)$ from $f \bullet g$ to a map into (F_b^1, F_b^0, e) . This proves the surjectivity.

Suppose $f_0, f_1 : (I, \partial I, 0) \rightarrow (F_b^1, F_b^0, e)$ are given, and let $K : (I, \partial I, 0) \times I \rightarrow (E_1, E_0, e)$ be a homotopy from f_0 to f_1 . We lift $\overline{p \circ K}$ to $L : I^2 \rightarrow E_0$ with initial condition $L(s, 0) = K(s, 1)$ and $L(0, t) = L(1, t) = e$. The homotopy $p \circ (K \bullet_2 L)$ is a homotopy of loops which is rel ∂I^2 homotopic to the constant map. We lift a homotopy to E_1 with initial condition $K \bullet_2 L$ on $I^2 \times \{0\}$ and constant on $\partial I^2 \times I$. The end is a homotopy from $f_0 \bullet \kappa_e$ to $f_1 \bullet \kappa_e$ (where κ_e is a constant loop at e). This proves the injectivity.

The higher dimensional case is obtained by an application to the relative Serre fibration $(\Omega^n F_b^1, \Omega^n F_b^0) \rightarrow (\Omega^n E_1, \Omega^n E_0) \rightarrow B$.

(2) Suppose $\pi_0(E) \rightarrow \pi_0(E_1)$ is surjective. The argument above for surjectivity is used to show the surjectivity of $\pi_0(F_b^0) \rightarrow \pi_0(F_b^1)$. The opposite implication is easy. ■

5.10 Higher connectivity

For many applications it is important to know that the homotopy groups of a space vanish in a certain range. We discuss several reformulations of this fact. In the following $\pi_0(X, x_0) = \pi_0(X)$. The space \mathbb{D}^0 is a singleton and $\mathbb{S}^{-1} = \emptyset$.

Proposition 5.10.1 ([26, Prop. 6.7.1]). *Let $n \geq 0$. TFAE:*

1. *$\pi_n(X, x) = 0$ for each $x \in X$.*
2. *Each map $\mathbb{S}^n \rightarrow X$ has an extension to \mathbb{D}^{n+1} .*
3. *Each map $\partial \mathbb{D}^{n+1} \rightarrow X$ has an extension to \mathbb{D}^{n+1} .*

Proof. The case $n = 0$ is trivial. The equivalence of (2) and (3) is a consequence of the homeomorphism $(\mathbb{D}^{n+1}, \mathbb{S}^n) \cong (I^{n+1}, \partial I^{n+1})$. Suppose $f : \mathbb{S}^n \rightarrow X$ is given. Use $e_1 = (1, 0, \dots) \in \mathbb{S}^n$ as a base point and think of f as representing an element of $\pi_n(X, x)$.

If (1) holds, then f is pointed null homotopic. A null homotopy $\mathbb{S}^n \times I \rightarrow X$ factors over the quotient map $\mathbb{S}^n \times I \rightarrow \mathbb{D}^{n+1}$, $(x, t) \mapsto (1-t)e_1 + tx$ and yields an extension of f . Conversely, let an element α of $\pi_n(X, x)$ be represented by a pointed map $f : (\mathbb{S}^n, e_1) \rightarrow (X, x)$. If this map has an extension F to \mathbb{D}^{n+1} , then (F, f) represents $\beta \in \pi_n(X, X, x) = 0$ with $\partial\beta = \alpha$. ■

Proposition 5.10.2 ([26, Prop. 6.7.3]). *Let $n \geq 1$. The following assertions about a topological pair (X, A) are equivalent:*

1. $\pi_n(X, A, *) = 0$ for each choice of $* \in A$.
2. Each map $f : (I^n, \partial I^n) \rightarrow (X, A)$ is as a map of pairs homotopic to a constant map.
3. Each map $f : (I^n, \partial I^n) \rightarrow (X, A)$ is homotopic rel ∂I^n to a map into A .

Proof. (1) \Rightarrow (2). Let $f : (I^n, \partial I^n) \rightarrow (X, A)$ be given. Since J^{n-1} is contractible, there exists a homotopy of the restriction $f : J^{n-1} \rightarrow A$ to a constant map. Since $J^{n-1} \subset \partial I^n$ and $\partial I^n \subset I^n$ are cofibrations, f is as a map of pairs homotopic to $g : (I^n, \partial I^n) \rightarrow (X, A)$ such that $g(J^{n-1}) = \{a_0\}$. Since $\pi_n(X, A, a_0) = 0$, the map $g : (I^n, \partial I^n, J^{n-1}) \rightarrow (X, A, a_0)$ is nullhomotopic as a map of triples.

(2) \Rightarrow (3) by the compression criterion (Proposition 5.2.3).

(3) \Rightarrow (1). Let $f : (I^n, \partial I^n, J^{n-1}) \rightarrow (X, A, *)$ be given. By assumption (3), $[f]$ is contained in the image of $\pi_n(A, A, *) \rightarrow \pi_n(X, A, *)$. Now use $\pi_n(A, A, *) = 0$. ■

Definition 5.10.3 (n -compressible pair). A topological pair (X, A) is called n -compressible if (1)-(3) in Proposition 5.10.2 hold. More generally, we call a map $f : X \rightarrow Y$ n -compressible if for each commutative diagram

$$\begin{array}{ccc} \partial \mathbb{S}^{n-1} & \xrightarrow{\varphi} & X \\ i \downarrow & \nearrow \Psi & \downarrow f \\ \mathbb{D}^n & \xrightarrow{\Phi} & Y \end{array} \quad (5.10.1)$$

there exists $\Psi : \mathbb{D}^n \rightarrow X$ such that $\Psi|_{\partial \mathbb{D}^n} = \varphi$ and $f \circ \Psi \simeq \Phi$ rel $\partial \mathbb{D}^n$ (this amounts to (3) of Proposition 5.10.2). This notion is invariant under compositions with homotopy equivalences of the target space Y .

Proposition 5.10.4 ([26, Prop. 6.7.5]). *Let $n \geq 0$. The following assertions about a pair (X, A) are equivalent:*

1. Each map $f : (\mathbb{D}^q, \partial \mathbb{D}^q) \rightarrow (X, A)$, $q \in (0, \dots, n)$ is rel $\partial \mathbb{D}^q$ homotopic to a map into A .

2. The inclusion $j : A \hookrightarrow X$ induces for each basepoint $a \in A$ a bijection $j_* : \pi_q(A, a) \rightarrow \pi_q(X, a)$ for $q < n$ and a surjection for $q = n$.
3. $\pi_q(X, A, a) = 0$ for $q = 1, \dots, n$ and each $a \in A$, and $\pi_0(A) \rightarrow \pi_0(X)$ is surjective.

Proof. (1) \Leftrightarrow (3). The surjectivity of $\pi_0(A) \rightarrow \pi_0(X)$ is equivalent to (1) for $q = 0$. The other cases follow from Proposition 5.10.2.

(2) \Leftrightarrow (3) follows from the exact sequence of the pair (Theorem 5.2.9). ■

Definition 5.10.5 (n -connected pair). A topological pair (X, A) is called n -connected if (1)-(3) in Proposition 5.10.4 hold. We call (X, A) ∞ -connected if the pair is n -connected for each n . A pair is ∞ -connected iff $j_* : \pi_n(A, a) \rightarrow \pi_n(X, a)$ is always bijective. If $X \neq \emptyset$ but $A = \emptyset$ we say that (X, A) is (-1) -connected, and (\emptyset, \emptyset) is ∞ -connected.

Definition 5.10.6 (Cone). Given a space X , its cone CX is defined as result of attaching the cylinder $X \times [0, 1]$ by its face $X \times \{0\}$ to a point $*$ along the projection $p : X \times \{0\} \rightarrow *$.

Proposition 5.10.7 ([26, Prop. 6.7.6]). *Let $n \geq 0$. The following assertions about a space X are equivalent:*

1. $\pi_q(X, x) = 0$ for $0 \leq q \leq n$ and $x \in X$.
2. The pair (CX, X) is $(n+1)$ -connected.
3. Each map $f : \partial I^q \rightarrow X$, $0 \leq q \leq n+1$ has an extension to I^q .

Proof. The cone CX is contractible, therefore $\partial : \pi_{q+1}(CX, X, *) \cong \pi_q(X, *)$. This and Proposition 5.10.4 shows the equivalence of (1) and (2). The equivalence of (1) and (3) is based on Proposition 5.10.1. ■

Definition 5.10.8 (n -connected space). A space X is called n -connected if (1)-(3) in Proposition 5.10.7 hold. Note that this is compatible with previous definitions for $n = 0$ (path-connected) and $n = 1$ (simply-connected).

Definition 5.10.9 (n -equivalence). Let $f : X \rightarrow Y$ be a map and $X \subset Z_f$ the inclusion into the mapping cylinder (recall the definition in Figure 5.1). Then f is said to be n -connected if (Z_f, X) is n -connected. We then also say that f is an n -equivalence. Thus f is n -connected iff $f_* : \pi_q(X, x) \rightarrow \pi_q(Y, f(x))$ is for each $x \in X$ bijective for $q < n$ and surjective for $q = n$. If f is an ∞ -equivalence we also call it a weak (homotopy) equivalence, which happens iff f_* is bijective for all $x \in X$ in all degrees.

Proposition 5.10.10 ([26, Prop. 6.7.7]). *Let $p : (E_1, E_0) \rightarrow B$ be a relative Serre fibration (see Proposition 5.9.14). Let (F_b^1, F_b^0) be the pair of fibers over $b \in B$. Then the pair (E_1, E_0) is n -connected iff the pairs (F_b^1, F_b^0) are n -connected for all $b \in B$.*

Proof. This is a direct consequence of Proposition 5.9.14. ■

Theorem 5.10.11 ([26, Theorem 6.7.9]). *Let $\varphi : (X, X_0, X_1) \rightarrow (Y, Y_0, Y_1)$ be a map such that the restrictions $\varphi_i : X_i \rightarrow Y_i$ are n -connected and $\varphi_{01} : X_0 \cap X_1 \rightarrow Y_0 \cap Y_1$ is $(n-1)$ -connected. Suppose $X = X_0^\circ \cup X_1^\circ$ and $Y = Y_0^\circ \cup Y_1^\circ$ where $^\circ$ denotes interiors. Then φ is an n -equivalence.*

Corollary 5.10.12 ([26, Prop. 6.7.10]). *Let $f : X \rightarrow Y$ be an n -connected map between spaces with nondegenerate basepoints. Then the suspension $\Sigma f : \Sigma X \rightarrow \Sigma Y$ is $(n+1)$ -connected. If X is n -connected, then ΣX is $(n+1)$ -connected. The sphere \mathbb{S}^{k+1} is k -connected.*

Proof. Let $\Sigma' X$ denote the unpointed suspension of X . This is a quotient of $X \times I$ and covered by the open cones $C_0 = X \times [0, 1]/X \times \{0\}$ and $C_1 = X \times (0, 1]/X \times \{1\}$ with intersection $X \times (0, 1)$. We can apply Proposition 5.10.10 directly; the cones are contractible and therefore the induced maps $C_j(X) \rightarrow C_j(Y)$ are ∞ -connected. In the case of a space X with a nondegenerate basepoint the quotient map $\Sigma' X \rightarrow \Sigma X$ is a homotopy equivalence. ■

Theorem 5.10.13 ([26, Theorem 6.7.11]). *Let $f : X \rightarrow Y$ be a continuous map. Let $\{U_j\}_j$ and $\{V_j\}_j$ be open coverings of X and Y such that $f(U_j) \subset V_j$. Suppose that for each finite subset of indices the induced map $f_J : \bigcap_{j \in J} U_j \rightarrow \bigcap_{j \in J} V_j$ is a weak equivalence. Then f is a weak equivalence.*

Proof. By passage to the mapping cylinder we can assume that f is an inclusion. Let $h : (I^n, \partial I^n) \rightarrow (Y, X)$ be given. We have to deform h rel ∂I^n into X . By compactness of I^n it suffices to work with finite coverings. A simple induction reduces the problem to a covering by only two sets $j = 0, 1$. Then we apply Theorem 5.10.11. ■

5.11 Excision for homotopy groups

Theorem 5.11.1 (Blakers-Massey excision theorem). *Let $X = U_0 \cup U_1$ be covered by two open subspaces U_0 and U_1 with non-empty intersection $V = U_0 \cap U_1$. Suppose that (U_0, V) is p -connected and (U_1, V) is q -connected, i.e.*

$$\begin{aligned}\pi_i(U_0, V, *) &= 0, & 1 \leq i \leq p, p \geq 0, \\ \pi_i(U_1, V, *) &= 0, & 1 \leq i \leq q, q \geq 0\end{aligned}\tag{5.11.1}$$

for each basepoint $$ $\in V$. Then the excision map, induced by the inclusion,*

$$i_* : \pi_n(U_1, V, *) \rightarrow \pi_n(X, U_0, *)\tag{5.11.2}$$

is an isomorphism for $1 \leq n < p + q$ and surjective for $n = p + q$ (for each $* \in X$). In the case $p = 0$, there is no condition on $\pi_i(U_0, V, *)$.

Corollary 5.11.2 ([26, Prop. 6.4.2]). *If in the above (U_1, V) is q -connected then (X, U_0) is q -connected.*

Proof. This is a special case of the excision theorem which also directly follows from Theorem 5.10.11. ■

For the sphere \mathbb{S}^n , introduce the following hemispheres and punctured spheres

$$\mathbb{D}_{\pm}^n = \{x \in \mathbb{S}^n \subset \mathbb{R}^{n+1} \mid \pm x_{n+1} \geq 0\} \subset H_{\pm}^n = \{x \in \mathbb{S}^n \mid x \neq \pm e_{n+1}\}. \quad (5.11.3)$$

We use the equatorial point $* = -e_1$ as a basepoint.

Lemma 5.11.3 ([26, Lemma 6.4.3]). *There are isomorphisms $\partial : \pi_{i+1}(D_-^{n+1}, \mathbb{S}^n, *) \rightarrow \pi_i(\mathbb{S}^n, *)$ for $i \geq 0, n \geq 0$ and $\pi_i(\mathbb{S}^n, *) \rightarrow \pi_i(\mathbb{S}^n, D_{\pm}^n, *)$ for $i \geq 0, n \geq 1$.*

Proof. In the first case we use the exact sequence of the pair $(\mathbb{D}_-^{n+1}, \mathbb{S}^n)$. The space \mathbb{D}_-^{n+1} is contractible and hence $\pi_i(\mathbb{D}_-^{n+1}, *) = 0$ for $i \geq 0$ and $n \geq 0$.

In the second case we consider similarly the exact sequence of $(\mathbb{S}^n, \mathbb{D}_{\pm}^n)$. Note that $* = -e_1 \in D_{\pm}^n$ for $n \geq 1$. ■

Thus for $n \geq 0$ we have a diagram with the isomorphisms

$$\begin{array}{ccc} \pi_i(\mathbb{S}^n, *) & \xrightarrow{E} & \pi_{i+1}(\mathbb{S}^{n+1}, *) \\ \uparrow \cong \partial & & \downarrow \cong \\ \pi_{i+1}(\mathbb{D}_-^{n+1}, \mathbb{S}^n, *) & \xrightarrow{i_*} & \pi_{i+1}(\mathbb{S}^{n+1}, D_{\pm}^n, *). \end{array} \quad (5.11.4)$$

The morphism i_* is induced by the inclusion and E is defined so as to make the diagram commutative. Note that the inductive proof of (1) in the next Theorem uses only Corollary 5.11.2.

Theorem 5.11.4 ([26, Thm 6.4.4]). *1. $\pi_i(\mathbb{S}^n) = 0$ for $i < n$ (recall that we already showed this in Corollary 5.6.3).*

- 2. The homomorphism i_* in diagram (5.11.4) is an isomorphism for $i \leq 2n - 2$ and an epimorphism for $i = 2n - 1$. A similar statement holds for E .*

Proof. Let $N(n)$ be the statement (1) and $E(n)$ be the statement (2). Obviously $N(1)$ holds. Assume $N(n)$ holds. We then deduce $E(n)$. We apply the excision theorem to $(X, U_0, U_1, V) = (\mathbb{S}^{n+1}, D_+^{n+1}, D_-^{n+1}, \mathbb{S}^n)$. By $N(n)$ and Lemma 5.11.3 we have $\pi_i(\mathbb{S}^n) \cong \pi_{i+1}(D_\pm^{n+1}, \mathbb{S}^n) = 0$ for $0 \leq i < n$. We use the excision theorem for $p = q = n$ and see that i_* is surjective for $i+1 \leq 2n$ and bijective for $i+1 \leq 2n-1$. Finally, $E(n)$ and $N(n)$ imply $N(n+1)$.

In order to have the correct hypotheses for the excision theorem, we thicken the spaces, replace D_\pm^n by H_\pm^n and note that the inclusions $D_\pm^n \subset H_\pm^n$ and $\mathbb{S}^{n-1} \subset H_+^n \cap H_-^n$ are homotopy equivalences. ■

Corollary 5.11.5 ([19, Prop. 5.11]). *The pair $(\mathbb{D}^{n+1}, \mathbb{S}^n)$ is n -connected.*

Proof. This follows from the exact sequence $\pi_r(\mathbb{D}^{n+1}) \rightarrow \pi_r(\mathbb{D}^{n+1}, \mathbb{S}^n) \rightarrow \pi_{r-1}(\mathbb{S}^n)$ for $r \leq n$, where the two groups on the sides vanish. ■

Proposition 5.11.6 ([26, Prop. 6.4.5]). *The homomorphism $\pi_i(\mathbb{D}^{n+1}, \mathbb{S}^n, *) \rightarrow \pi_i(\mathbb{D}^{n+1}/\mathbb{S}^n, *)$ induced by the quotient map is an isomorphism for $i \leq 2n-1$ and an epimorphism for $i = 2n$.*

Proof. Consider the commutative diagram

$$\begin{array}{ccc} \pi_i(D_-^{n+1}, \mathbb{S}^n, *) & \longrightarrow & \pi_i(D_-^{n+1}/\mathbb{S}^n, *) \\ i_* \downarrow & & \downarrow (1) \\ \pi_{i+1}(\mathbb{S}^{n+1}, \mathbb{D}_+^{n+1}, *) & \xrightarrow{(2)} & \pi_i(\mathbb{S}^{n+1}/D_+^{n+1}, *). \end{array} \quad (5.11.5)$$

The map (1) is induced by a homeomorphism and the map (2) by a homotopy equivalence, hence both are isomorphisms. Now apply Theorem 5.11.4. ■

The homomorphism E is essentially the suspension homomorphism:

$$\Sigma_* : \pi_n(X) \xleftarrow[\cong]{\partial} \pi_{n+1}(CX, X) \xrightarrow{q_*} \pi_{n+1}(CX/X) = \pi_{n+1}(\Sigma X) \quad (5.11.6)$$

with the quotient map $q : \mathbb{D}^{n+1} \rightarrow \mathbb{S}^{n+1} = \mathbb{D}^{n+1}/\mathbb{S}^n$. The next result is the famous suspension theorem of Freudenthal.

Theorem 5.11.7 (Freudenthal's suspension theorem [26, Theorem 6.4.6]). *The suspension $\Sigma_* : \pi_i(\mathbb{S}^n) \rightarrow \pi_{i+1}(\mathbb{S}^{n+1})$ is an isomorphism for $i \leq 2n - 2$ and an epimorphism for $i = 2n - 1$.*

Proof. We have to show that $q_* : \pi_{i+1}(CX, X) \rightarrow \pi_{i+1}(CX/X)$ is for $X = \mathbb{S}^n$ an isomorphism (epimorphism) in the appropriate range. This follows from Proposition 5.11.6; one has to use that \mathbb{D}_-^{n+1} is the (pointed) cone on \mathbb{S}^n . ■

Theorem 5.11.8 ([26, Theorem 6.4.7]). *$\pi_n(\mathbb{S}^n) \cong \mathbb{Z}$ and $\Sigma_* : \pi_n(\mathbb{S}^n) \rightarrow \pi_{n+1}(\mathbb{S}^{n+1})$ is an isomorphism ($n \geq 1$). The group $\pi_n(\mathbb{S}^n)$ is generated by the identity map of \mathbb{S}^n .*

Proof. From the exact sequence of the Hopf fibration $\pi_2(\mathbb{S}^3) \rightarrow \pi_2(\mathbb{S}^2) \rightarrow \pi_1(\mathbb{S}^1) \rightarrow \pi_1(\mathbb{S}^3)$ and $\pi_j(\mathbb{S}^3) = 0$ for $j = 1, 2$, we obtain an isomorphism $\partial : \pi_2(\mathbb{S}^2) \rightarrow \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$. From Theorem 5.11.7 we obtain a surjection $\Sigma_* : \pi_1(\mathbb{S}^1) \rightarrow \pi_2(\mathbb{S}^2)$; this is an isomorphism, since both groups are isomorphic to \mathbb{Z} . For $n \geq 2$, Theorem 5.11.7 directly gives an isomorphism Σ_* . We know that $\pi_1(\mathbb{S}^1)$ is generated by the identity map, and Σ_* respects the identity. ■

This proof clearly extends to two cones over any space X , so by using the adjunction $[\Sigma X, Y] \cong [X, \Omega Y]$ we get the more general form of this Theorem.

Theorem 5.11.9 (Freudenthal's suspension theorem). *Let X be an n -connected space, i.e. $\pi_i(X) = 0$ for $0 \leq i \leq n$. Then the homomorphism $\pi_k(X) \rightarrow \pi_k(\Omega\Sigma X)$ induced by the inclusion $X \hookrightarrow \Omega\Sigma X$ is an isomorphism for $k \leq 2n - 2$ and an epimorphism for $k = 2n - 1$. Furthermore, since $\pi_{k+1}(\Sigma X) \cong [\mathbb{S}^{k+1}; \Sigma X] \cong [\mathbb{S}^k; \Omega\Sigma X] \cong \pi_k(\Omega\Sigma X)$, for $k \leq 2n - 2$ there is an isomorphism $\pi_k(X) \cong \pi_{k+1}(\Sigma X)$.*

Proof. See [26, Theorem 6.4.6]. ■

Definition 5.11.10 (Degree on spheres). For a map $f : \mathbb{S}^n \rightarrow \mathbb{S}^n$, its degree $\deg f$ is the integer such that $[f] = [\text{id}_{\mathbb{S}^n}]^{\deg f}$ in $\pi_n(\mathbb{S}^n)$.

Example 5.11.11 (Hopf fibrations). We continue the discussion of the Hopf fibrations 5.9.12. The Hopf fibration $\mathbb{S}^{2n+1} \rightarrow \mathbb{C}P^n$ and $\pi_i(\mathbb{S}^{2n+1}) = 0$ for $i \leq 2n$ yield $\pi_2(\mathbb{C}P^n) \cong \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$ and $\pi_i(\mathbb{C}P^n) = 0$ for $0 \leq i \leq 2n$, $i \neq 2$. The inclusion $\mathbb{S}^{2n+1} \rightarrow \mathbb{S}^{2n+3}$, $z \mapsto (z, 0)$ induces an embedding $\mathbb{C}P^n \subset \mathbb{C}P^{n+1}$. We compare the corresponding Hopf fibrations and their exact sequences and conclude $\pi_2(\mathbb{C}P^n) \cong \pi_2(\mathbb{C}P^{n+1})$. Let $\mathbb{C}P^\infty = \bigcup_{n \geq 1} \mathbb{C}P^n$ be the colimit (the topology on it is defined as the finest topology under which all inclusions $\mathbb{C}P^n \hookrightarrow \mathbb{C}P^\infty$ are continuous). The inclusions induce $\pi_i(\mathbb{C}P^n) \cong \pi_i(\mathbb{C}P^\infty)$ for $i \leq 2n$. A proof uses the fact that a compact subset of $\mathbb{C}P^\infty$ is contained in some finite $\mathbb{C}P^N$. Therefore $\mathbb{C}P^\infty$ is a space with a single nontrivial homotopy group $\pi_2(\mathbb{C}P^\infty) \cong \mathbb{Z}$.

Particularly notable is the special case

$$\pi_3(\mathbb{S}^2) \cong \pi_3(\mathbb{S}^3) \cong \mathbb{Z}. \quad (5.11.7)$$

This is the first of the “nontrivial” homotopy groups of spheres $\pi_k(\mathbb{S}^n)$, $k > n$, and the only one that can be computed without much more advanced homology theory.

We have the similar results for real projective spaces. The double coverings $\mathbb{S}^n \rightarrow \mathbb{R}P^n$ are used to show that $\pi_1(\mathbb{R}P^2) \cong \pi_1(\mathbb{R}P^3) \cong \dots \cong \pi_1(\mathbb{R}P^\infty) \cong \mathbb{Z}_2$, induced by the inclusions $\mathbb{R}P^n \hookrightarrow \mathbb{R}P^{n+1}$ for $i < n$ and $\pi_i(\mathbb{R}P^n) = 0$ for $0 \leq i < n$, $i \neq 1$. The space $\mathbb{R}P^\infty$ therefore has a single nontrivial homotopy group $\pi_1(\mathbb{R}P^\infty) \cong \mathbb{Z}_2$.

Example 5.11.12 (Lens spaces). Recall the lens space $L = L_m(l_1, \dots, l_n)$ from Example 5.7.12. It can be given a structure of a CW-complex with one cell in each dimension from 0 to $2n - 1$ so that the degree of the attaching maps alternates between 0 and m (starting with degree 0 for the map $\partial\mathbb{D}^1 \rightarrow L^{(0)}$).

Example 5.11.13 ($\pi_3(\mathbb{S}^2 \vee \mathbb{S}^2) \cong \mathbb{Z}^3$). Let us compute $\pi_3(\mathbb{S}^2 \vee \mathbb{S}^2)$. We have an inclusion $i : \mathbb{S}^2 \vee \mathbb{S}^2 \hookrightarrow \mathbb{S}^2 \times \mathbb{S}^2$, because $\mathbb{S}^2 \times \mathbb{S}^2$ can be realized as a CW-complex that differs from $\mathbb{S}^2 \vee \mathbb{S}^2$ only by an addition of a 4-cell. Observe that the induced homomorphism on the homotopy groups is surjective:

$$i_* : \pi_n(\mathbb{S}^2 \vee \mathbb{S}^2) \rightarrow \pi_n(\mathbb{S}^2 \times \mathbb{S}^2) = \pi_n(\mathbb{S}^2) \oplus \pi_n(\mathbb{S}^2). \quad (5.11.8)$$

(Here we used the fact that a smash product with \mathbb{S}^k is homeomorphic to the iterated reduced suspension Σ^k , and $\Sigma^k \mathbb{S}^n \cong \mathbb{S}^{n+k}$.) This allows us to insert extra zeroes into the long exact sequence for the pair $(\mathbb{S}^2 \times \mathbb{S}^2, \mathbb{S}^2 \vee \mathbb{S}^2)$ without breaking its exactness, which splits it into the short exact sequences

$$0 \rightarrow \pi_{n+1}(\mathbb{S}^2 \times \mathbb{S}^2, \mathbb{S}^2 \vee \mathbb{S}^2) \rightarrow \pi_n(\mathbb{S}^2 \vee \mathbb{S}^2) \xrightarrow{i_*} \pi_n(\mathbb{S}^2 \times \mathbb{S}^2) \rightarrow 0. \quad (5.11.9)$$

For $n = 3$, since $\mathbb{S}^2 \vee \mathbb{S}^2$ is 1-connected and the remaining 4-cell is 4-connected, we can use the excision theorem to compute

$$\pi_4(\mathbb{S}^2 \times \mathbb{S}^2, \mathbb{S}^2 \vee \mathbb{S}^2) \cong \pi_4(\mathbb{S}^2 \times \mathbb{S}^2 / \mathbb{S}^2 \vee \mathbb{S}^2) \cong \pi_4(\mathbb{S}^2 \wedge \mathbb{S}^2) \cong \pi_4(\mathbb{S}^4) \cong \mathbb{Z}. \quad (5.11.10)$$

At the same time, $\pi_3(\mathbb{S}^2 \times \mathbb{S}^2) = \pi_3(\mathbb{S}^2) \oplus \pi_3(\mathbb{S}^2) = \mathbb{Z}^2$ (generated by the Hopf fibrations from the preceding Example 5.11.11). Thus we have the exact sequence

$$0 \rightarrow \mathbb{Z} \rightarrow \pi_3(\mathbb{S}^2 \vee \mathbb{S}^2) \rightarrow \mathbb{Z}^2 \rightarrow 0, \quad (5.11.11)$$

which implies $\pi_3(\mathbb{S}^2 \vee \mathbb{S}^2) = \mathbb{Z}^3$ because every short exact sequence of the form $0 \rightarrow A \xrightarrow{i} B \xrightarrow{\rho} \mathbb{Z}^2 \rightarrow 0$ splits (the splitting monomorphism $j : \mathbb{Z}^2 \rightarrow B$ can be constructed by defining

it on the basis $j(1, 0) = b_1$ and $j(0, 1) = b_2$, where $b_1, b_2 \in B$ are any two elements such that $p(b_1) = (1, 0)$ and $p(b_2) = (0, 1)$; it is then easy to check that $p \circ j = \text{id}_{\mathbb{Z}^2}$ and hence $B \cong A \oplus \mathbb{Z}^2$.

This construction is a special case of the *Whitehead product* $\pi_k(X) \times \pi_l(X) \rightarrow \pi_{k+l-1}(X)$. Given $f : \mathbb{S}^k \rightarrow X$ and $g : \mathbb{S}^l \rightarrow X$, their Whitehead product is the homotopy class denoted

$$[f, g] \in \pi_{k+l-1}(X) \quad (5.11.12)$$

and represented by the composition

$$\mathbb{S}^{k+l-1} \xrightarrow{\chi} \mathbb{S}^k \vee \mathbb{S}^l \xrightarrow{f \vee g} X, \quad (5.11.13)$$

where χ is the attaching map for the single $(k+l)$ -cell in the CW-complex $\mathbb{S}^k \times \mathbb{S}^l$ which consists of $\mathbb{S}^k \vee \mathbb{S}^l$ and that cell (χ maps the boundary of the $(k+l)$ -disk to the $(k+l-1)$ -skeleton, which in this case is $\mathbb{S}^k \vee \mathbb{S}^l$). In the case above, the third “unexpected” generator of $\pi_3(\mathbb{S}^2 \vee \mathbb{S}^2)$ is $[i_1, i_2]$, where $i_1, i_2 : \mathbb{S}^2 \rightarrow \mathbb{S}^2 \vee \mathbb{S}^2$ are the inclusion maps for the two 2-spheres.

Another simple Whitehead product is $[\text{id}_{\mathbb{S}^2}, \text{id}_{\mathbb{S}^2}] \in \pi_3(\mathbb{S}^2)$. It is represented by the map obtained by gluing $\mathbb{S}^2 \vee \mathbb{S}^2 \subset \mathbb{S}^3$ onto \mathbb{S}^2 . Since we already know that $\pi_3(\mathbb{S}^2)$ is generated by the Hopf fibration p_{Hopf} and this map clearly has winding number 2, we find that

$$[\text{id}_{\mathbb{S}^2}, \text{id}_{\mathbb{S}^2}] = 2[p_{\text{Hopf}}] \quad (5.11.14)$$

(although there might also be a minus sign depending on how one defines the Hopf fibration).

Exercise 5.4 ([26, Thm. 6.10.5]). Let X and Y be well-pointed (i.e. with non-degenerate basepoints) and assume that $\pi_i(X) = \pi_j(Y) = 0$ for $i < p, j < q$ with $p, q \geq 2$. Show that the inclusion $X \vee Y \hookrightarrow X \times Y$ induces an isomorphism of the π_i -groups for $i \leq p+q-2$. Also show that $\pi_i(X \times Y, X \vee Y)$ and $\pi_i(X \wedge Y)$ vanish for $i \leq p+q-1$.

Exercise 5.5 ([26, Exercise 6.11]). Extend the above example to determine $\pi_{2n-1}(\mathbb{S}^n \vee \mathbb{S}^n)$ for $n \geq 2$.

Part II

Differential Geometry

6 Lie Theory I: Basics

6.1 Lie groups

Definition 6.1.1 (Topological group). A topological group is a group G together with a topology on it that is compatible with the group structure. That is, both the inversion $g \mapsto g^{-1}$ and the multiplication $(g, h) \mapsto gh$ maps are continuous. This is equivalent to requiring that the map $G \times G \rightarrow G$ given by $(g, h) \mapsto gh^{-1}$ is continuous.

Topological groups form a very big class and include things like “infinite-dimensional Lie groups”, p -adic numbers, rational numbers, etc. It turns out that every locally compact and locally contractible topological group (in particular topological groups that are also topological manifolds) have a unique smooth structure compatible with the group structure, which is what a Lie group is. Hence we skip right to the definition of Lie groups, and for the most part will assume that they are smooth.

Definition 6.1.2 (Lie group). A Lie group is a group G with a structure of a differentiable (or C^k -differentiable, C^∞ -smooth, C^ω -analytic) manifold on it that is compatible with the group structure. That is, the map $G \times G \rightarrow G$ given by $(g, h) \mapsto gh^{-1}$ is differentiable (resp. C^k , C^∞ , C^ω).

We will show that any differentiable Lie group has a unique compatible real analytic structure. In fact, one can show that differentiability is not required at all, and any topological manifold with a continuous group structure has a unique structure of an analytic Lie group. This is why most authors assume that Lie groups are smooth by definition.

Proposition 6.1.3 ([33, Thm. 1.6.3]). *Any twice-differentiable Lie group G has a unique structure of a real analytic manifold which turns it into a real analytic Lie group, and such that the identity map on G is a C^2 -diffeomorphism between the two structures.*

Definition 6.1.4 (Group translation and adjoint operators). The left and right translation operators on a group G are defined by

$$L_g(h) = gh, \quad R_g(h) = hg. \quad (6.1.1)$$

The adjoint operator corresponding to an element $g \in G$ is

$$\mathbf{Ad}_g h = L_g \circ R_g^{-1}(h) = ghg^{-1}. \quad (6.1.2)$$

Example 6.1.5 (Lie groups). (a) Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, or \mathbb{H} (quaternions, see below). Every \mathbb{K} -vector space V is a Lie group with respect to its additive structure and the manifold structure of the real vector space obtained from V by restriction of the algebraic structure. The dimension is $\dim_{\mathbb{R}} \mathbb{K} \cdot \dim V$.

- (b) Every finite or countable group with discrete topology is a Lie group.
- (c) The direct product $G \times H$ of two Lie groups G, H is a Lie group with the product group structure and the differentiable structure of the product of manifolds.
- (d) $\mathrm{GL}_n(\mathbb{R})$ is the set of invertible $n \times n$ matrices with real entries. It is a group under matrix multiplication and an open subset of $\mathrm{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Matrix multiplication is obviously smooth and inversion is smooth by Cramer's rule.
- (e) $\mathrm{GL}_n^+(\mathbb{R})$, which is the subset of $\mathrm{GL}_n(\mathbb{R})$ consisting of matrices with positive determinant. It is still a group and an open subset of $\mathrm{GL}_n(\mathbb{R})$.
- (f) If $H \subset G$ is an open subgroup (a subgroup that is an open subset) of a Lie group G , then H is a Lie group with the inherited group and manifold structures.
- (g) For any real or complex finite-dimensional vector space V , the group $\mathrm{GL}(V)$ of endomorphisms of V is a Lie group.
- (h) The unit circle $\mathbb{S}^1 \subset \mathbb{C}^*$ with multiplication inherited from \mathbb{C} is a Lie group.
- (i) The n -torus $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ is an n -dimensional abelian Lie group.

Definition 6.1.6 (Lie group homomorphism). A Lie group homomorphism between two Lie groups G, H is a smooth map $F : G \rightarrow H$ that is also a group homomorphism. If it is a diffeomorphism, it is called an isomorphism. These maps comprise the morphisms of the category LieGr of Lie groups.

- Example 6.1.7** (Lie group morphisms).
- (a) The inclusion map $\mathbb{S}^1 \hookrightarrow \mathbb{C}^*$ is a Lie group homomorphism.
 - (b) The exponential map $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_+, *)$ is a Lie group homomorphism since $\exp(x+y) = \exp(x) \cdot \exp(y)$.
 - (c) The determinant function $\det : \mathrm{GL}_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$ is a Lie group homomorphism.
 - (d) The adjoint action $\mathrm{Ad}_g : G \rightarrow G$ is a Lie group homomorphism for any $g \in G$.

Theorem 6.1.8 ([7, Thm. 7.5]). *Every Lie group homomorphism has constant rank.*

Proof. Let $F : G \rightarrow H$ be a Lie group homomorphism. We will show that the differential $T_g F$ of F at any $g \in G$ has the same rank as the differential $T_e F$ at the identity. We have

$$F \circ L_g(h) = F(gh) = F(g)F(h) = L_{F(g)} \circ F(h), \quad (6.1.3)$$

i.e. $F \circ L_g = L_{F(g)} \circ F$. Computing the differential of both sides at the identity $e \in G$ using the chain rule, we find

$$T_g F \circ T_e L_g = T_{F(g)} L_{F(g)} \circ T_e F. \quad (6.1.4)$$

Left translation operators are diffeomorphisms and therefore both $T_e L_g$ and $T_{F(g)} L_{F(g)}$ are isomorphisms of tangent spaces. Compositions with isomorphisms don't change rank, therefore $\text{rank } T_g F = \text{rank } T_e F$ is constant on G . ■

Corollary 6.1.9 ([16, Cor. 5.3.7]). *For a Lie group homomorphism $\varphi : G \rightarrow H$, the following are equivalent:*

1. φ is an immersion (submersion).
2. φ has discrete kernel (is open).
3. φ_{*e} is injective (surjective).

In addition, φ is an isomorphism of Lie groups iff it is bijective.

Proof. Exercise 1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1. ■

The simplest nonabelian examples of compact Lie groups are the 3-dimensional rotation group and its “complex form”, SU_2 . We now establish that SU_2 is the universal covering space of $SO_3(\mathbb{R})$, and hence $\pi_1(SO_3(\mathbb{R})) = \mathbb{Z}_2$. As we will show in the next section, every Lie group has a unique universal covering Lie group.

Example 6.1.10 ($SO_3(\mathbb{R})$ and its double cover SU_2). The rotation group $SO_3(\mathbb{R})$ is known to be a Lie group because, for example, it can be parametrized by Euler angles as coordinates in a way that makes the group structure obviously smooth. We will have a general proof that orthogonal groups are Lie groups in Section 6.4. This group can be alternatively parametrized by vectors $n \in \mathbb{R}^3$ as follows:

$$R_n = I + \frac{\sin |n|}{|n|} A_n + \frac{1 - \cos |n|}{|n|} A_n^2, \quad (6.1.5)$$

where $|n| = \sqrt{n_1^2 + n_2^2 + n_3^2}$ is the rotation angle and

$$A_\omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \quad A_\xi = \begin{pmatrix} -\omega_3^2 - \omega_2^2 & \omega_1 \omega_2 & \omega_3 \omega_1 \\ \omega_1 \omega_2 & -\omega_3^2 - \omega_1^2 & \omega_2 \omega_3 \\ \omega_3 \omega_1 & \omega_2 \omega_3 & -\omega_1^2 - \omega_2^2 \end{pmatrix}. \quad (6.1.6)$$

It is easy to check two important equalities

$$A_{\omega \times \xi} = A_\omega A_\xi - A_\xi A_\omega = [A_\omega, A_\xi], \quad \omega \cdot \xi = -\text{tr}(A_\omega A_\xi). \quad (6.1.7)$$

If $|n| \in 2\pi\mathbb{Z}$, then R_n is a rotation around the axis through n , through the angle $|n|$. Furthermore, if $0 < |n| < \pi$, then R_n is the unique rotation such that, for each nonzero vector $x \perp n$ in the plane of the rotation, the triple $(x, R_n \cdot x, n)$ has positive orientation,

i.e. $\det(x, R_n x, n) > 0$. The mapping $n \mapsto R_n$ is surjective from the closed ball $|n| \leq \pi$ onto $\mathrm{SO}_3(\mathbb{R})$, and a diffeomorphism from the open ball $|n| < \pi$ onto the dense subset of $\mathrm{SO}_3(\mathbb{R})$ of rotations through an angle $< \pi$. Near the boundary sphere $S_\pi = \{n \in \mathbb{R}^3 \mid |n| = \pi\}$, one can check that $n \mapsto R_n$ is a smooth two-fold covering map mapping n and $(|n| - 2\pi) \frac{n}{|n|}$ to the same element of $\mathrm{SO}_3(\mathbb{R})$. In particular, $n \rightarrow R_n$ is a smooth two-fold covering from the sphere S_π onto the subset of rotations through the angle π , mapping the antipodal points $\pm n$ to the same element of $\mathrm{SO}_3(\mathbb{R})$. We thus see that $\mathrm{SO}_3(\mathbb{R})$ is diffeomorphic to the real projective space \mathbb{RP}^3 , which is the quotient of the closed ball \mathbb{D}^3 by the map that identifies antipodal boundary points, or equivalently the antipodal quotient of \mathbb{S}^3 .

The group SU_2 is defined as the set of 2×2 complex matrices that preserve the standard Hermitian dot product $(z, w)^2 = |z|^2 + |w|^2$ on \mathbb{C}^2 :

$$\mathrm{SU}_2 = \left\{ g(a, b) = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \middle| a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1 \right\}. \quad (6.1.8)$$

Clearly this is diffeomorphic to the unit sphere in $\mathbb{C}^2 \cong \mathbb{R}^4$, i.e. $\mathrm{SU}_2 \cong \mathbb{S}^3$ as manifolds. One remarkable property is that

$$\mathbb{R} \cdot \mathrm{SU}_2 = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \middle| a, b \in \mathbb{C} \right\} \quad (6.1.9)$$

is a 4-dimensional real-linear subspace of the (real 8-dimensional) space of complex 2×2 matrices. This subspace is closed under matrix multiplication and in this way forms a noncommutative real division algebra. It has the famous quaternion basis consisting of the identity matrix **1** plus the three matrices

$$\mathbf{i} = i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{k} = i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (6.1.10)$$

with relations $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -\mathbf{1}$. So $\mathbb{R} \cdot \mathrm{SU}_2$ can be identified with the quaternion algebra \mathbb{H} and SU_2 is the multiplicative group of unit quaternions.

6.2 Topology of Lie groups

Recall that not every topological space has a universal covering space, and in particular not every topological group has one. However Lie groups are manifolds and therefore always have a universal covering manifold. We now show that the universal cover naturally becomes a Lie group itself.

Theorem 6.2.1 (Existence of a universal covering group [7, Thm. 7.7]). *Let G be a connected Lie group. There exists a simply connected Lie group \tilde{G} , called the universal covering group of G , that admits a smooth covering map $\tilde{\pi} : \tilde{G} \rightarrow G$ that is also a Lie group homomorphism. This group is unique up to Lie group isomorphism.*

Proof. Let \tilde{G} be the universal covering manifold of G and $\tilde{\pi}$ the corresponding smooth covering map. Then $\tilde{\pi} \times \tilde{\pi} : \tilde{G} \times \tilde{G} \rightarrow G \times G$ is also a smooth covering map.

Let $m : G \times G \rightarrow G$ and $i : G \times G$ denote the multiplication and inversion maps of G , respectively, and let \tilde{e} be an arbitrary element of the fiber $\tilde{\pi}^{-1}(e) \subset \tilde{G}$. Since \tilde{G} is simply connected, the Lifting Theorem 4.1.18 guarantees that the map $m \circ (\tilde{\pi} \times \tilde{\pi}) : \tilde{G} \times \tilde{G} \rightarrow G$ has a unique continuous lifting $\tilde{m} : \tilde{G} \times \tilde{G} \rightarrow \tilde{G}$ satisfying $\tilde{m}(\tilde{e}, \tilde{e}) = \tilde{e}$ and $\tilde{\pi} \circ \tilde{m} = m \circ (\tilde{\pi} \times \tilde{\pi})$:

$$\begin{array}{ccc} \tilde{G} \times \tilde{G} & \xrightarrow{\tilde{m}} & \tilde{G} \\ \tilde{\pi} \times \tilde{\pi} \downarrow & & \downarrow \tilde{\pi} \\ G \times G & \xrightarrow{m} & G. \end{array} \quad (6.2.1)$$

Because $\tilde{\pi}$ is a local diffeomorphism and $\tilde{\pi} \circ \tilde{m} = m \circ (\tilde{\pi} \times \tilde{\pi})$ is smooth, it follows that \tilde{m} is smooth. By the same reasoning, $i \circ \tilde{\pi} : \tilde{G} \rightarrow G$ has a smooth lifting $\tilde{i} : \tilde{G} \rightarrow \tilde{G}$ satisfying $\tilde{i}(\tilde{e}) = \tilde{e}$ and $\tilde{\pi} \circ \tilde{i} = i \circ \tilde{\pi}$:

$$\begin{array}{ccc} \tilde{G} & \xrightarrow{\tilde{i}} & \tilde{G} \\ \tilde{\pi} \downarrow & & \downarrow \tilde{\pi} \\ G & \xrightarrow{i} & G. \end{array} \quad (6.2.2)$$

We define multiplication and inversion on \tilde{G} by $xy = \tilde{m}(x, y)$ and $x^{-1} = \tilde{i}(x)$. Then the diagrams above can be written as

$$\tilde{\pi}(xy) = \tilde{\pi}(x)\tilde{\pi}(y), \quad \tilde{\pi}(x^{-1}) = \tilde{\pi}(x)^{-1}. \quad (6.2.3)$$

It remains only to show that \tilde{G} is a group with these operations, for then it is a Lie group because \tilde{m} and \tilde{i} are smooth, and then the last formula shows that $\tilde{\pi}$ is a homomorphism. We leave the details, which can be found in [7, Thm. 7.7], to the reader, since we will later provide another proof as a consequence of Theorem 11.3.16.

Uniqueness follows from the uniqueness of covering spaces up to isomorphism and the fact that the only flexibility in the definition of the group structure is in the choice of the identity \tilde{e} , but all such choices lead to isomorphic groups. ■

Corollary 6.2.2. *For a connected Lie group we have the exact sequence*

$$0 \rightarrow \pi_1(G) \rightarrow \tilde{G} \rightarrow G \rightarrow 0. \quad (6.2.4)$$

- Example 6.2.3** (Universal covering groups). (a) \mathbb{R}^n is the universal covering group of the torus \mathbb{T}^n .
(b) \mathbb{C} is the universal covering group of \mathbb{C}^* with the covering map $\exp : \mathbb{C} \rightarrow \mathbb{C}^*$.

Proposition 6.2.4. *For any topological group G , the fundamental group $\pi_1(G)$ is abelian.*

Proof. Fundamentally this follows from the fact that there are two different multiplication operations on loops in a group – concatenation of loops, and multiplications of their values in the group. Specifically, given two loops γ_1, γ_2 , a map $g(s, t) = \gamma_1(s)\gamma_2(t)$ provides a homotopy between $\gamma_1 \bullet \gamma_2$ (along two sides of the square) and $\gamma_2 \bullet \gamma_1$ (along the other two sides). ■

We have already seen that $\pi_1(\mathrm{SO}_3(\mathbb{R})) = \mathbb{Z}_2$ by finding its universal covering group. Unfortunately, computing the fundamental groups of other classical groups is not so easy and will require considering the long exact sequences of certain fibrations, which will be easier to construct once we talk about group actions on manifolds. Another nontrivial fact from Lie theory will be that the fundamental group of any compact Lie group is finite, which means that the universal covering group of any compact Lie group is also compact.

First recall the following fundamental theorem that classifies all finitely generated abelian groups (it has nothing to do with the theory of Lie groups, so the proof can be skipped).

Theorem 6.2.5 (Classification of finitely generated abelian groups). *Every finitely generated abelian group is isomorphic to a group of the form*

$$\mathbb{Z}^n \oplus \mathbb{Z}_{k_1} \oplus \cdots \oplus \mathbb{Z}_{k_t}. \quad (6.2.5)$$

Proof. First we prove the following statement: if the abelian group G is generated by a set x_1, \dots, x_k , and $c_1, \dots, c_k \in \mathbb{N}$ are such that $\gcd(c_1, \dots, c_k) = 1$, then there exists another set of generators $y_1, \dots, y_k \in G$ such that $y_1 = c_1x_1 + \cdots + c_kx_k$.

We argue by induction on $s = c_1 + \cdots + c_k$. The lemma certainly holds if $s = 1$, and so we assume $s > 1$. Then at least two c_i are nonzero, say $c_1 \geq c_2 \geq 0$. Now

- $\{x_1, x_2 + x_1, x_3, \dots, x_k\}$ generates G ,
- $\gcd(c_1 - c_2, c_2, c_3, \dots, c_k) = 1$, and
- $(c_1 - c_2) + c_2 + \cdots + c_k < s$,

and so, by induction, there exist generators y_1, \dots, y_k for G such that

$$y_1 = (c_1 - c_2)x_1 + c_2(x_1 + x_2) + c_2x_3 + \cdots + c_kx_k = c_1x_1 + \cdots + c_kx_k. \quad (6.2.6)$$

Now we prove the classification theorem by induction on k . If G can be generated by one element, the statement is trivial, so we may assume that $k > 1$. Among all generating sets $\{x_1, \dots, x_k\}$ for G with k elements there is one for which the order of x_1 is the smallest possible. We shall show that G is then the direct sum of $\langle x_1 \rangle$ and $\langle x_2, \dots, x_k \rangle$. This will complete the proof because the induction hypothesis provides us with a basis for the second group, which together with x_1 forms a basis for G .

If G is not the direct sum of $\langle x_1 \rangle$ and $\langle x_2, \dots, x_k \rangle$, then there exists a relation

$$m_1x_1 + m_2x_2 + \cdots + m_kx_k = 0 \quad (6.2.7)$$

with $m_1x_1 \neq 0$. After possibly changing the sign of some of the x_i , we may suppose that $m_1, \dots, m_k \in \mathbb{N}$ and $m_1 < \text{ord}(x_1)$. Let $d = \gcd(m_1, \dots, m_k) > 0$ and let $c_i = m_i/d$. According to the first part of the proof, there exists a generating set y_1, \dots, y_k such that $y_1 = c_1x_1 + \cdots + c_kx_k$. But

$$dy_1 = m_1x_1 + \cdots + m_kx_k = 0 \quad (6.2.8)$$

and $d \leq m_1 < \text{ord}(x_1)$, and so this contradicts the choice of $\{x_1, \dots, x_k\}$. ■

Remark 6.2.6. The *torsion* subgroup of G is defined as the subset of all elements of finite order:

$$T(G) = \{x \in G : \exists m \in \mathbb{N} \text{ } mx = 0\}, \quad (6.2.9)$$

and then the theorem implies that each finitely generated abelian group has the form

$$G \cong \mathbb{Z}^n \oplus T(G) \quad (6.2.10)$$

where $T(G)$ is a finite abelian group.

It is important to know that the power n of \mathbb{Z} is uniquely determined by G and is called the *rank* of G . Indeed, pick a prime p not dividing any of the k_i 's and consider the quotient

$$G/pG \cong (\mathbb{Z}/p\mathbb{Z})^n, \quad (6.2.11)$$

therefore n is the dimension of G/pG viewed as a vector space over the finite field \mathbb{F}_p , and thus depends only on G itself.

Furthermore, k_i 's can be chosen so that $k_1 \geq 2$ and $k_1|n_2, \dots, k_{t-1}|k_t$, and then they are also uniquely determined by G (such k_i are called *invariant factors* of G). Alternatively, k_i 's can be chosen to be powers of prime numbers, which also determines them uniquely (such k_i

are called elementary divisors of G). Both of these statements follow from the fact that if $\gcd(k, l) = 1$ then $\mathbb{Z}_k \times \mathbb{Z}_l \cong \mathbb{Z}_{kl}$, which allows one to decompose the torsion part of G into a product of cyclic groups of prime power orders.

Definition 6.2.7 (Centralizer). Let G be a group and $S \subset G$ a subset. The centralizer of S is the subgroup

$$C(S) = \{g \in G \mid gsg^{-1} = s \text{ for all } s \in S\}. \quad (6.2.12)$$

The center of G is defined as $Z(G) = C(G)$. If all elements of S commute with each other, then $C(S)$ is the largest subgroup of G whose center contains S .

The following properties hold. G is abelian iff $Z(G) = G$. A subgroup H is abelian iff $H \subset C(H)$. There is an inclusion $C(S) \subset N(S)$, and in fact $C(S)$ is a normal subgroup of $N(S)$.

Proposition 6.2.8. Suppose G is a connected topological group and $\Gamma \subset G$ is a discrete normal subgroup. Then $\Gamma \subset Z(G)$.

Proof. Let $\gamma \in \Gamma$ and consider the continuous map $g \mapsto \mathbf{Ad}_g(\gamma) = g\gamma g^{-1}$ for $g \in G$. Since Γ is normal, $\mathbf{Ad}_G(\gamma) \subset \Gamma$. Since G is connected, the map is continuous, and Γ is discrete, the image $\mathbf{Ad}_G(\gamma)$ must be a single point, and it must be γ since we can choose $g = e$. ■

This gives us yet another way of proving that $\pi_1(G)$ is always abelian.

Corollary 6.2.9. Every connected Lie group G has the form \tilde{G}/Γ , where \tilde{G} is simply connected and $\Gamma \cong \pi_1(G)$ is a discrete subgroup of the center $Z(\tilde{G})$. In particular, $\pi_1(G)$ is abelian.

Corollary 6.2.10 (Classification of discrete subgroups of \mathbb{R}^n). Any discrete subgroup Γ of \mathbb{R}^n is a free abelian subgroup, i.e. a lattice generated by a set of linearly independent vectors x_1, \dots, x_k .

Proof. Indeed, consider the subspace $V \subset \mathbb{R}^n$ spanned by Γ . Let its dimension be k and let $v_1, \dots, v_k \in \Gamma$ be a set of elements of Γ that forms a basis for V , and let $Z = \{\mathbb{Z}v_1 + \dots + \mathbb{Z}v_k \cong \mathbb{Z}^k\} \cong \mathbb{Z}^k$. Consider the quotient map $q : V \rightarrow V/Z$ (this is a quotient by a subgroup, not just a topological quotient: its fibers are $q^{-1}(v) = v + Z$). It is open and the quotient space is compact (being homeomorphic to a torus \mathbb{T}^k). The fact that Γ is discrete and q -saturated (i.e. $q^{-1}(q(\Gamma)) = \Gamma$) implies that $q(\Gamma)$ is discrete.

Moreover, Γ is closed because otherwise there would exist a Cauchy sequence $(\gamma_k) \in \Gamma$, i.e. $\gamma_k - \gamma_{k+1} \rightarrow 0$ (the existence of a limit is irrelevant for this proof, a mere existence of a Cauchy sequence is enough), which would contradict the fact that 0 is isolated in Γ . Therefore, by definition of the quotient topology and since Γ is q -saturated, $q(\Gamma)$ is also

closed. Therefore $q(\Gamma) \cong \Gamma/\mathbb{Z}^k$ is compact and discrete, hence finite. Since \mathbb{Z}^k and Γ/\mathbb{Z}^k are both finitely generated, it follows that Γ is also finitely generated (by the union of the generators of \mathbb{Z}^k and Γ/\mathbb{Z}^k). Since V is torsion-free, we obtain that Γ is free abelian. The fact that Γ/\mathbb{Z}^k is finite then implies that $\text{rank } \Gamma = k$. Therefore there exists a set of generators x_1, \dots, x_k for Γ which also provide a basis for V . ■

Remark 6.2.11. We will later show that \mathbb{R}^n are the only simply connected abelian Lie groups, and it will follow that any other connected abelian Lie group G is a quotient of \mathbb{R}^n by a lattice of rank $k \leq n$, so $G \cong \mathbb{T}^k \oplus \mathbb{R}^{n-k}$.

This gives us another way of proving the Fundamental Theorem of Algebra. If \mathbb{F}/\mathbb{R} is a field extension, then the identity component $(\mathbb{F}^*)_0$ is abelian and connected. Then for extension of degree $d = 1$, this is \mathbb{R} ; for $d = 2$ this is $\mathbb{S}^1 \times \mathbb{R} \cong \mathbb{C}^*$, and for $d \geq 3$ this is $\mathbb{S}^{d-1} \times \mathbb{R}$, but none of these can be abelian Lie groups. Hence \mathbb{C} is the largest extension of \mathbb{R} and hence algebraically closed.

6.3 Lie subgroups

Definition 6.3.1 (Lie subgroup). A Lie subgroup H of a Lie group G is an immersed submanifold such that the induced group structure turns it into a Lie group. In particular, the inclusion map $i : H \hookrightarrow G$ is a smooth immersion and a Lie group homomorphism.

Note that Lie subgroups are not generally required to be embedded submanifolds. For example, the dense immersion of \mathbb{R} into the torus \mathbb{T}^2 (as a straight line with an irrational slope) is a Lie homomorphism. We will show later that its image is a Lie subgroup.

Proposition 6.3.2 ([7, Prop. 7.11]). *Let G be a Lie group, and suppose $H \subset G$ is a subgroup that is also an embedded submanifold. Then H is a Lie subgroup.*

Proof. We only need to check that the multiplication and inversion maps on H are smooth. Both of them are restrictions of smooth maps either $G \times G \rightarrow G$ or $G \rightarrow G$ to $H \times H \rightarrow G$ and $H \rightarrow G$, and are therefore smooth. Since H is an embedded submanifold, restricting the codomain of a smooth map into G to H also produces a smooth map by Theorem 3.9.12. ■

Lemma 6.3.3 ([7, Lem. 7.12]). *Let $H \subset G$ be an open subgroup of the Lie group G . Then H is an embedded Lie subgroup. In addition, H is closed, so it is a union of connected components of G .*

Proof. Any open subset is an embedded submanifold. Every left coset gH is open in G since it is an image of an open subset under the diffeomorphism L_g . Since the complement $G \setminus H$ is the union of the cosets of H other than H , it is open, and therefore H is closed. ■

The following is a simple exercise.

Proposition 6.3.4 ([7, Prop. 7.15]). *Let G be a Lie group and let G_0 be the connected component of the identity $e \in G$. Then G_0 is a normal subgroup of G , and is the only connected open subgroup. Every connected component of G is diffeomorphic to G_0 .*

Definition 6.3.5 (Identity component). The identity component of a Lie group G is the connected component G_0 of the identity $e \in G$.

Proposition 6.3.6 ([7, Prop. 7.14]). *The identity component $G_0 \subset G$ is generated by any connected open neighborhood U of $e \in G$ (i.e. it is the smallest subgroup containing U).*

Proof. Let H be the subgroup generated by U . WLOG, we can take U to be a symmetric neighborhood of the identity, i.e. $U^{-1} = U$ (otherwise simply replace it with $U \cap U^{-1}$, which is also open). Let U^k be the set of all elements of G that can be expressed as a product of k elements of U . Then $H = \bigcup_{k=1}^{\infty} U^k$ (this is a simple exercise). Since multiplication is a diffeomorphism, each U^k can be written as a union of open sets, and thus H is open.

Since U is connected and multiplication is continuous, each U^k is connected. Thus H is also connected by virtue of being a union of overlapping (at e) connected sets.

By Lemma 6.3.3, H must be the identity component. ■

Proposition 6.3.7 ([7, Prop. 7.16]). *Let $F : G \rightarrow H$ be a Lie group homomorphism. Then $\ker F$ is a properly embedded Lie subgroup of G of codimension rank F .*

Proof. Because F has constant rank, its kernel $F^{-1}(e)$ is a properly embedded submanifold of codimension equal to rank F by Theorem 3.9.7. It is thus a Lie subgroup by Proposition 6.3.2. ■

Proposition 6.3.8 ([7, Prop. 7.17]). *If $F : G \rightarrow H$ is an injective Lie group homomorphism, the image of F has a unique smooth manifold structure such that $F(G)$ is a Lie subgroup of H and $F : G \rightarrow F(G)$ is a Lie group isomorphism.*

Proof. Since a Lie group homomorphism has constant rank, it follows by the Global Rank Theorem 3.7.5 that F is a smooth immersion. Proposition 3.9.3 shows that $F(G)$ has a unique smooth structure such that it is an immersed submanifold of H and F is a diffeomorphism onto its image. It is a Lie group (because G is), and it is a subgroup for algebraic reasons, therefore it is a Lie subgroup. Because $F : G \rightarrow F(G)$ is a group isomorphism and a diffeomorphism, it is a Lie group isomorphism. ■

Remark 6.3.9. As we will show later, a similar theorem holds for the image of *any* Lie group homomorphism, not just injective.

Example 6.3.10 (Dense Lie subgroup of the torus). Let $H \subset \mathbb{T}^2$ be the dense submanifold of the torus that is the image of the immersion $\gamma : \mathbb{R} \rightarrow \mathbb{T}^2$ defining a straight line with an irrational slope. It is easy to check that γ is an injective Lie group homomorphism, and thus H is an immersed Lie subgroup of \mathbb{T}^2 . Note that $\gamma(\mathbb{R})$ is neither an open nor a closed subgroup of \mathbb{T}^2 . Indeed, as we will show in the Closed Subgroup Theorem, any closed subgroup of a Lie group has to be an embedded Lie subgroup.

Example 6.3.11 (Embedded Lie Subgroups). (a) $\mathrm{SL}_n(\mathbb{R})$ is the kernel of the determinant homomorphism $\det : \mathrm{GL}_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$, and is therefore a properly embedded Lie subgroup. Since \det is surjective, it is a smooth submersion by the global rank theorem, so $\mathrm{SL}_n(\mathbb{R})$ has dimension $n^2 - 1$.

(b) Similarly $\mathrm{SL}_n(\mathbb{C})$ is a properly embedded Lie subgroup of $\mathrm{GL}_n(\mathbb{C})$ of real dimension $2n^2 - 2$.

Typically we will be interested in closed subgroups, and the following result is a very important property.

Theorem 6.3.12 ([7, Thm. 7.21]). *A Lie subgroup of a Lie group is closed iff it is embedded.*

We skip the proof of this theorem since later on we will derive a much stronger version of it, the Closed Subgroup Theorem 12.4.2, which asserts that *every* closed subgroup (not necessarily a submanifold!) of a Lie group is automatically embedded.

6.4 Classical Lie groups

Example 6.4.1 (Classical groups). Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. For $n, m = 0, 1, 2, \dots$, define

$$I_{n,m} = \begin{bmatrix} I_n & 0 \\ 0 & -I_m \end{bmatrix}, \quad J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \quad (6.4.1)$$

A classical Lie group can be most generally defined as the subgroup of $\mathrm{GL}(\mathbb{K}^n)$ consisting of transformations that preserve some nondegenerate bilinear or sesquilinear form $B : \mathbb{K}^2 \rightarrow \mathbb{K}$. However, there are no nontrivial bilinear forms over \mathbb{H} , and only \mathbb{R} admits symmetric bilinear forms with a signature (which isn't all pluses). Hermitian forms have basis-independent signature in both the complex and the quaternionic case (the real case reduces to the symmetric one). A skew-Hermitian form on a complex vector space is rendered Hermitian by multiplication by i , so in this case, only \mathbb{H} is interesting. Thus, with some work, one can show that any classical Lie group is isomorphic to one of the groups we define here (the [Wikipedia page](#) has a full derivation).

1. Real and complex unimodular groups

$$\mathrm{SL}_n(\mathbb{K}) = \{a \in \mathrm{GL}_n(\mathbb{K}) \mid \det a = 1\}, \mathbb{K} = \mathbb{R}, \mathbb{C}. \quad (6.4.2)$$

2. The real orthogonal group

$$\mathrm{O}_{n,m} = \{a \in \mathrm{GL}_{n+m}(\mathbb{R}) \mid a^T I_{n,m} a = I_{n,m}\}. \quad (6.4.3)$$

In the case $m = 0$ we write O_n or $\mathrm{O}_n(\mathbb{R})$ for the group of linear isometries of the Euclidean scalar product.

The special orthogonal group is the subgroup of $\mathrm{O}_{n,m}$ of isometries that preserve orientation, i.e. the unimodular ones

$$\mathrm{SO}_{n,m} = \mathrm{O}_{n,m} \cap \mathrm{SL}_{n+m}(\mathbb{R}). \quad (6.4.4)$$

In the case $m = 0$ we write SO_n or $\mathrm{SO}_n(\mathbb{R})$.

3. The complex orthogonal group

$$\mathrm{O}_n(\mathbb{C}) = \{a \in \mathrm{GL}_n(\mathbb{C}) \mid a^T a = I_n\}. \quad (6.4.5)$$

The special complex orthogonal group is the subgroup of $\mathrm{O}_n(\mathbb{C})$ of unimodular isometries

$$\mathrm{SO}_n(\mathbb{C}) = \mathrm{O}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C}). \quad (6.4.6)$$

4. The unitary group

$$\mathrm{U}_{n,m} = \{a \in \mathrm{GL}(n+m, \mathbb{C}) \mid a^\dagger I_{n,m} a = I_{n,m}\}. \quad (6.4.7)$$

In the case $m = 0$ we write U_n for the group of linear isometries of the standard Hermitian scalar product on \mathbb{C}^n .

The special unitary group is the subgroup of unimodular unitary transformations

$$\mathrm{SU}_{n,m} = \mathrm{U}_{n,m} \cap \mathrm{SL}_{n+m}(\mathbb{C}). \quad (6.4.8)$$

In the case $m = 0$ we write SU_n .

5. The symplectic group is the group of isometries of the antisymmetric bilinear form $\omega(x, y) = (x, J_n y)$ on \mathbb{K}^{2n} ,

$$\mathrm{Sp}_n(\mathbb{K}) = \{a \in \mathrm{GL}_{2n}(\mathbb{K}) \mid a^T J_n a = J_n\}. \quad (6.4.9)$$

6. The quaternionic symplectic group is the group of isometries of the Hermitian form $(\bar{x}, I_{n,m} y)$ on \mathbb{H}^{n+m} ,

$$\mathrm{Sp}_{n,m} = \{a \in \mathrm{GL}_{n,m}(\mathbb{H}) \mid a^\dagger I_{n,m} a = I_{n,m}\}. \quad (6.4.10)$$

In the case $m = 0$ we write Sp_n for the group of linear isometries of the standard scalar product on \mathbb{H}^n . Note that these groups are in fact different from their complex and real versions. The difference is explained in detail in [37, pp. 5-6].

Since we already know that $\mathrm{GL}_n(\mathbb{K})$ is a smooth manifold, we can use the Level Set Theorem 3.9.7 to conclude that all of these classical groups are also smooth properly embedded manifolds. The multiplication and inversion operations on them are smooth by virtue of being the restrictions of the ones on $\mathrm{GL}_n(\mathbb{K})$ and hence these groups are Lie groups. The groups O_n , U_n , and Sp_n also turn out to be compact because, for example, all of their elements have bounded matrix norms (e.g. real orthogonal matrices of size $n \times n$ have matrix norm \sqrt{n}).

Example 6.4.2 ($SU_2 \cong \mathrm{Sp}_1$). Recall that the quaternion algebra \mathbb{H} is isomorphic to the algebra of 2×2 complex matrices, $\mathrm{Mat}_2(\mathbb{C})$:

$$a_0 I + a_1 i + a_2 j + a_3 k \mapsto \begin{pmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{pmatrix}, \quad a_0, a_1, a_2, a_3 \in \mathbb{R}. \quad (6.4.11)$$

It is easy to check that the image of Sp_1 under this homomorphism coincides with SU_2 . Hence, by restriction in range, it induces a group isomorphism from Sp_1 onto SU_2 which is smooth in both directions and hence a Lie group isomorphism.

Example 6.4.3 (Universal coverings of SO_3 and SO_4). Identify \mathbb{R}^4 with \mathbb{H} by the isometric isomorphism of real vector spaces

$$\lambda(\mathbf{x}) = x_0 \mathbf{1} + x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k}. \quad (6.4.12)$$

Every pair of quaternions (\mathbf{a}, \mathbf{b}) defines a real-linear mapping $\mathbb{H} \rightarrow \mathbb{H}$ by $q \mapsto \mathbf{a}q\bar{\mathbf{b}}$ (where $\mathbf{b} = b_0 \mathbf{1} - b_1 \mathbf{i} - b_2 \mathbf{j} - b_3 \mathbf{k}$) and hence a linear mapping $\phi(\mathbf{a}, \mathbf{b})$ of \mathbb{R}^4 by

$$\phi(\mathbf{a}, \mathbf{b})x = \lambda^{-1}(\mathbf{a}\lambda(x)\bar{\mathbf{b}}), \quad x \in \mathbb{R}^4. \quad (6.4.13)$$

By assigning $\phi(\mathbf{a}, \mathbf{b})$ to (\mathbf{a}, \mathbf{b}) one obtains a map $\mathbb{H} \oplus \mathbb{H} \rightarrow \mathrm{Mat}_4(\mathbb{R})$. This is a homomorphism of real algebras, and in particular it is smooth.

First consider $\phi(\mathbf{a}, \mathbf{a})$ for $a \in \mathrm{Sp}_1$. Due to $\|\mathbf{a}\|^2 = 1$, $\phi(\mathbf{a}, \mathbf{a})$ is orthogonal. Since it leaves invariant the subspace $\mathbb{R} \times \{0\} \subset \mathbb{R}^4$, it also leaves invariant the subspace $\{0\} \times \mathbb{R}^3 \subset \mathbb{R}^4$. Hence, restriction of ϕ in domain to the submanifold $\{(\mathbf{a}, \mathbf{a}) \mid \mathbf{a} \in \mathrm{Sp}_1\}$ of $\mathbb{H} \oplus \mathbb{H}$ and in range to the embedded submanifold of $\mathrm{Mat}_4(\mathbb{R})$ consisting of the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}, \quad R \in O_3, \quad (6.4.14)$$

yields a Lie group homomorphism $\phi : \mathrm{Sp}_1 \rightarrow O_3$. The defining equation reduces to

$$\phi(\mathbf{a})x = \lambda^{-1}(\mathbf{a}(x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k})\bar{\mathbf{a}}), \quad x \in \mathbb{R}^3. \quad (6.4.15)$$

For $\mathbf{a} \in \mathbb{H}$, an explicit calculation yields

$$[\phi(\mathbf{a})]_{ij} = (a_0^2 - a_1^2 - a_2^2 - a_3^2)\delta_{ij} + 2(a_i a_j - a_0 \epsilon_{ijk} a_k). \quad (6.4.16)$$

We determine the kernel and the image of ϕ . According to (6.4.15), $\mathbf{a} \in \ker \phi$ iff \mathbf{a} commutes with all quaternions. Then $\mathbf{a} = a\mathbf{1}$, $a \in \mathbb{R}$, hence $\ker \phi = \{\pm \mathbf{1}\}$, the center of Sp_1 . To find $\mathrm{im} \phi$, one first shows that ϕ is an immersion (exercise). Then, since Sp_1 and SO_3 have the same dimension, ϕ is a submersion and hence an open mapping by Proposition 3.10.4. It follows that $\mathrm{im} \phi$ contains a neighborhood of the identity $I \in \mathrm{SO}_3$. Since $\mathrm{im} \phi$ is a subgroup, by Proposition 6.3.6, it contains the identity component SO_3 of O_3 . Since Sp_1 is connected, we finally obtain $\mathrm{im} \phi = \mathrm{SO}_3$.

Since Sp_1 coincides with the unit sphere in \mathbb{H} and since $\ker \phi = \{\pm \mathbf{1}\}$ implies that the preimage of a point in SO_3 under ϕ consists of antipodal points, ϕ induces a bijection $\mathbb{RP}^3 \cong \mathrm{SO}_3$. Since both ϕ and the projection $\mathbb{S}^3 \rightarrow \mathbb{RP}^3$ are submersions, this bijection is in fact a diffeomorphism.

Next, consider $\phi(\mathbf{a}, \mathbf{b})$ for $\mathbf{a}, \mathbf{b} \in \mathrm{Sp}_1$. Due to $\|\mathbf{a}\|^2 = \|\mathbf{b}\|^2 = 1$, $\phi(\mathbf{a}, \mathbf{b})$ is orthogonal. Hence, by restriction in domain and range, ϕ induces a Lie group homomorphism $\phi : \mathrm{Sp}_1 \times \mathrm{Sp}_1 \rightarrow \mathrm{O}_4$ with the defining equation (6.4.13). Arguing as before one finds $\ker \phi = \{\pm (\mathbf{1}, \mathbf{1})\}$ and $\mathrm{im} \phi = \mathrm{SO}_4$. In particular, we conclude that there is a homeomorphism $\mathrm{SO}_4 \cong \mathbb{S}^3 \times \mathrm{SO}_3$.

Let us add that via the algebra homomorphism (6.4.11), all of the above has an equivalent formulation in terms of complex 2×2 matrices. This formulation is obtained by replacing Sp_1 by SU_2 , \mathbb{H} by the real vector space spanned by I_2 and the traceless skew-Hermitian matrices, equipped with the scalar product $(A, B) = \mathrm{tr}(A^\dagger B)$, and the subspace of \mathbb{H} spanned by $\mathbf{i}, \mathbf{j}, \mathbf{k}$ by the subspace of traceless skew-Hermitian matrices. This exactly recovers our previous construction of SU_2 as a double cover of $\mathrm{SO}_3(\mathbb{R})$.

Remark 6.4.4. Using octonions one can construct in the same way a homeomorphism $\mathrm{SO}_8 \cong \mathbb{S}^7 \times \mathrm{SO}_7$. But in all other cases SO_n is only a ‘twisted product’ of SO_{n-1} and \mathbb{S}^{n-1} , see [22, Cor. 4D.3].

Example 6.4.5 (Universal covering of $\mathrm{SO}_{1,3}^+$). Let $S_2(\mathbb{C})$ be the 4-dimensional real vector space of 2×2 complex Hermitian matrices parametrized by the mapping

$$\lambda : \mathbb{R}^4 \rightarrow S_2(\mathbb{C}), \quad \lambda(x) = \begin{pmatrix} x_0 + x_1 & -ix_2 + x_3 \\ ix_2 + x_3 & x_0 - x_1 \end{pmatrix}. \quad (6.4.17)$$

It is isometric with respect to the quadratic forms $x_0^2 - x_1^2 - x_2^2 - x_3^2$ on \mathbb{R}^4 and $A \mapsto \det A$ on $S_2(\mathbb{C})$. For every $a \in \mathrm{Mat}_2(\mathbb{C})$, the assignment $A \mapsto aAa^\dagger$ defines a linear mapping of $S_2(\mathbb{C})$ and hence a linear mapping $\phi(a)$ of \mathbb{R}^4 given by

$$\phi(a)x = \lambda^{-1}(a\lambda(x)a^\dagger). \quad (6.4.18)$$

The assignment $a \mapsto \phi(a)$, in turn, defines a map $\phi : \text{Mat}_2(\mathbb{C}) \rightarrow \text{Mat}_4(\mathbb{R})$. This map is real homogenous of degree 2, hence smppth, and satisfies $\phi(ab) = \phi(a)\phi(b)$. Since for every $a \in \text{SL}_2(\mathbb{C})$ we have $\det a A a^\dagger = \det A$, ϕ restricts to a Lie group homomorphism $\phi : \text{SL}_2(\mathbb{C}) \rightarrow \text{O}_{1,3}$. By similar arguments as in the last example, one finds $\ker \phi = \{\pm I\}$ (which is the center of $\text{SL}_2(\mathbb{C})$), $\text{im } \phi = \text{SO}_{1,3}^+$ (the identity component of $\text{SO}_{1,3}$, which is the *proper orthochronous Lorentz group* that preserves both spatial and temporal orientations) and that ϕ is a covering homomorphism. It is a universal covering because $\text{SL}_2(\mathbb{C})$ is simply connected. To see this, one may use the polar decomposition which yields a diffeomorphism $\text{U}_2 \times S_2(\mathbb{C}) \rightarrow \text{GL}_2(\mathbb{C})$, given by $(a, A) \mapsto ae^A$. Restriction of this diffeomorphism to the submanifold SU_2 of U_2 times the subspace of traceless elements of $S_2(\mathbb{C})$ induces a diffeomorphism $\mathbb{S}^3 \times \mathbb{R}^3 \rightarrow \text{SL}_2(\mathbb{C})$. Hence, ϕ is the universal covering homomorphism and $\text{SL}_2(\mathbb{C})$ is the universal covering group of $\text{SO}_{1,3}^+$.

As we have just discovered, the universal covering group of SO_3 is $\text{SU}_2 \cong \text{Sp}_1$. It is also called the spin group Spin_3 . However, not every universal covering group of a classical group is a classical group itself. The higher-dimensional spin groups, for instance, are not isomorphic to classical groups except for a few “accidental isomorphisms”.

Definition 6.4.6 (Pin and Spin groups). For $n \geq 3$, the universal covering group of O_n is called Pin_n , and that of SO_n is called Spin_n .

We note that this is not the “proper” definition of spin groups, which is actually a larger class that also includes $\text{Pin}_{n,m}$, $\text{Spin}_{n,m}$, and their complex versions, none of which are generally simply connected and hence are not universal covers of any classical groups.

Proposition 6.4.7. *The following deformation retracts exist:*

$$\text{GL}_n(\mathbb{R}) \rightarrow \text{O}_n(\mathbb{R}), \quad \text{GL}_n(\mathbb{C}) \rightarrow \text{U}_n, \quad \text{GL}_n(\mathbb{H}) \rightarrow \text{Sp}_n, \quad (6.4.19)$$

$$\text{SL}_n(\mathbb{R}) \rightarrow \text{SO}_n(\mathbb{R}), \quad \text{SL}_n(\mathbb{C}) \rightarrow \text{SU}_n. \quad (6.4.20)$$

Proof. This is done by Gram-Schmidt orthonormalization and its complex and quaternionic analogs. To any ordered basis (v_1, v_2, \dots, v_n) of the space $V = \mathbb{K}^n$ associate the “flag” of subspaces $V_0 = \{0\}$, $V_1 = \langle v_1 \rangle$, $V_2 = \langle v_1, v_2 \rangle$, and so on until $V_n = V$. The Gram-Schmidt algorithm turns any such basis into an orthonormal basis (b_1, \dots, b_n) that gives rise to the same flag of subspaces. It is moreover the unique such basis (orthonormal and with this same flag) for which in addition each b_i , inside V_i , is on the same side of the hyperplane V_{i-1} as the original basis vector v_i .

Now taking $V = \mathbb{R}^n$ we can identify $\text{GL}_n^+(\mathbb{R})$ (the identity component of $\text{GL}_n(\mathbb{R})$) with the set of ordered bases $\mathbf{v} = (v_1, \dots, v_n)$ with $\det \mathbf{v} > 0$, and $\text{SO}_n(\mathbb{R})$ with the set of orthonormal bases \mathbf{b} with $\det \mathbf{b} = 1$. Now for such a basis \mathbf{v} let \mathbf{b} be the corresponding Gram-Schmidt

orthonormal basis, and simultaneously (or successively) deform every v_i linearly to b_i , as $t \mapsto (1-t)v_i + tb_i$. The intermediate vectors stay inside V_i , and since b_i is on the same side as v_i , they never enter V_{i-1} . This means the deformed vectors stay linearly independent at all times, so the deformation takes place inside $\mathrm{GL}_n(\mathbb{R})$. Since we started in $\mathrm{GL}_n^+(\mathbb{R})$, this means that we end in it as well, and in particular in $\mathrm{SO}_n(\mathbb{R})$. This is in fact a strong deformation retract: elements of $\mathrm{SO}_n(\mathbb{R})$ stay fixed the whole time. The second component of $\mathrm{GL}_n(\mathbb{R})$ is diffeomorphic to the identity component and thus retracts in the same way onto the second component of $\mathrm{O}_n(\mathbb{R})$.

The same approach works for \mathbb{C} and \mathbb{H} , though it might be easier to rewrite the Gram-Schmidt process in terms of matrix factorization. For the unimodular groups, notice that $\mathrm{GL}_n(\mathbb{K}) \cong \mathbb{K}^* \times \mathrm{SL}_n(\mathbb{K})$ and the restrictions of the above deformation retractions to $\{1\} \times \mathrm{SL}_n(\mathbb{K})$ give exactly the stated groups. ■

This is a special case of a very important structural theorem from Lie theory: every Lie group deformation retracts into a maximal compact subgroup (which is unique up to conjugation). Therefore the homotopy types of Lie groups are completely captured by compact Lie groups, which in turn are completely classified. In turn, as we will see, this means that finite-dimensional representation theory of Lie groups can be reduced to the compact case.

Exercise 6.1. Write $Z(G)$ for the center of a Lie group G . Show that:

- (a) $Z(\mathrm{U}_n) \cong \mathrm{U}_1$ and $Z(\mathrm{SU}_n) \cong \mathbb{Z}_n$ for $n \geq 2$,
- (b) $Z(\mathrm{O}_{2n}) \cong \mathbb{Z}_2$, $Z(\mathrm{SO}_{2n}) \cong \mathbb{Z}_2$ for $n \geq 2$, and $Z(\mathrm{SO}_2) = \mathrm{SO}_2$,
- (c) $Z(\mathrm{SO}_{2n+1}) \cong \mathbb{Z}_2$ for $n \geq 1$ and $Z(\mathrm{SO}_{2n+1}) \cong 1$ for $n \geq 1$,
- (d) $Z(\mathrm{Sp}_n) \cong \mathbb{Z}_2$.

Exercise 6.2 (Iwasawa decomposition). (a) Let $A \subset \mathrm{GL}_n(\mathbb{R})$ be the subgroup of diagonal matrices with positive elements on the diagonal and let $N \subset \mathrm{GL}_n(\mathbb{R})$ be the subgroup of upper singular matrices with 1's on the diagonal. Using Gram-Schmidt orthogonalization, show multiplication induces a diffeomorphism of $\mathrm{O}_n \times A \times N$ onto $\mathrm{GL}_n(\mathbb{R})$. This is called the *Iwasawa*, or *KAN*, decomposition of $\mathrm{GL}_n(\mathbb{R})$. As topological spaces, show that $\mathrm{GL}_n(\mathbb{R}) \cong \mathrm{O}_n \times \mathbb{R}^{n(n+1)/2}$. Similarly, as topological spaces, show that $\mathrm{SL}_n(\mathbb{R}) \cong \mathrm{SO}_n \times \mathbb{R}^{(n+2)(n+1)/2}$.

- (b) Let $A \subset \mathrm{GL}_n(\mathbb{C})$ be the subgroup of diagonal matrices with positive real elements on the diagonal and let $N \subset \mathrm{GL}_n(\mathbb{C})$ be the subgroup of upper singular matrices with 1's on the diagonal. Show multiplication induces a diffeomorphism of $\mathrm{U}_n \times A \times N$ onto $\mathrm{GL}_n(\mathbb{C})$. As topological spaces, show that $\mathrm{GL}_n(\mathbb{C}) \cong \mathrm{U}_n \times \mathbb{R}^{n^2}$. Similarly, as topological spaces, show that $\mathrm{SL}_n(\mathbb{R}) \cong \mathrm{SU}_n \times \mathbb{R}^{n^2-1}$.

Exercise 6.3 (Bruhat decomposition). Let $N \subset \mathrm{GL}_n(\mathbb{C})$ be the subgroup of upper triangular matrices with 1's on the diagonal, let $\bar{N} \subset \mathrm{GL}_n(\mathbb{C})$ be the subgroup of lower triangular matrices

with 1's on the diagonal, and let W be the subgroup of permutation matrices (i.e. matrices with a single one in each row and each column and zeros elsewhere). Use Gaussian elimination to show $\mathrm{GL}_n(\mathbb{C}) = \bigsqcup_{w \in W} \bar{N}wN$. This is called the *Bruhat decomposition* for $\mathrm{GL}_n(\mathbb{C})$.

- Exercise 6.4.** (a) Let $P \subset \mathrm{GL}_n(\mathbb{R})$ be the set of positive definite symmetric matrices. Show that multiplication gives a bijection from $P \times \mathrm{O}_n$ to $\mathrm{GL}_n(\mathbb{R})$.
(b) Let $H \subset \mathrm{GL}_n(\mathbb{C})$ be the set of positive definite Hermitian matrices. Show that multiplication gives a bijection from $H \times \mathrm{U}_n$ to $\mathrm{GL}_n(\mathbb{C})$.

Exercise 6.5. Complex general linear, real orthogonal, unitary, and real symplectic groups are closely related in a way that is fundamental to much of differential geometry and physics. Namely, the following intersections hold (where all groups are viewed as subsets of $\mathrm{GL}_{2n}(\mathbb{R})$):

$$\mathrm{U}_n = \mathrm{O}_{2n} \cap \mathrm{Sp}_{2n}(\mathbb{R}) = \mathrm{O}_{2n} \cap \mathrm{GL}_n(\mathbb{C}) = \mathrm{Sp}_{2n}(\mathbb{R}) \cap \mathrm{GL}_n(\mathbb{C}). \quad (6.4.21)$$

In other words, the unitary group coincides with the intersection of any two out of three groups O_{2n} , $\mathrm{Sp}_{2n}(\mathbb{R})$, and $\mathrm{GL}_n(\mathbb{C})$. In differential geometry this leads to the result that having any two out of three structures – Riemannian metric, symplectic form, or a complex structure – naturally leads to a compatible third structure. This underlies, for instance, the definition of Kähler geometry. Prove the above equalities. *Hint:* e.g. for the first intersection, show that if a transformation preserves the Euclidean bilinear form $g(x, y) = (x, I_{2n}y)$ and the symplectic form $\omega(x, y) = (x, J_{2n}y)$, then it also preserves the form $h(u, v) = g(u, v) + i\omega(u, v)$, which defines a Hermitian dot product if we let the first n coordinates be the real parts, the last n coordinates the imaginary parts, and define the structure of a complex vector space by putting $iu = Ju$. Similar constructions also exist for the other two directions.

6.5 CW-structure of classical groups

The CW-structure of SO_n will come with a very nice cellular map $\rho : \mathbb{R}P^{n-1} \times \mathbb{R}P^{n-2} \times \cdots \times \mathbb{R}P^1 \rightarrow \mathrm{SO}_n$. To simplify notation we will write P^i for $\mathbb{R}P^i$.

To each nonzero $v \in \mathbb{R}^m$ we can associate the reflection $r(v) \in \mathrm{O}_n$ across the hyperplane v^\perp . Since $r(v)$ is a reflection, it has determinant -1 , so to get an element of SO_n we consider $\rho(v) = r(v)r(e_1)$ where $e_1 = (1, 0, \dots, 0)$. Since $\rho(v)$ depends only on the line spanned by v , it defines a map $P^{n-1} \rightarrow \mathrm{SO}_n$. This map is injective since it is the composition of $v \mapsto r(v)$, which is obviously an injection of P^{n-1} into $\mathrm{O}_n \setminus \mathrm{SO}_n$, with the homeomorphism $\mathrm{O}_n \setminus \mathrm{SO}_n \rightarrow \mathrm{SO}_n$ given by the right multiplication by $r(e_1)$. Since ρ is injective and P^{n-1} is compact Hausdorff, we may think of ρ as embedding P^{n-1} as a subspace of SO_n .

More generally, for a sequence $I = (i_1, \dots, i_m)$ with each $i_j < n$, we define a map $\rho : P^I = P^{i_1} \times \cdots \times P^{i_m} \rightarrow \mathrm{SO}_n$ by

$$\rho(v_1, \dots, v_m) = \rho(v_1) \cdots \rho(v_m). \quad (6.5.1)$$

If $\varphi^i : \mathbb{D}^i \rightarrow P^i$ is the standard characteristic map for the i -cell of P^i , restricting to the 2-sheeted covering projection $\partial\mathbb{D}^i \rightarrow P^{i-1}$ (see Example 5.3.17), then the product $\varphi^I : \mathbb{D}^I \rightarrow P^I$ of the appropriate φ^{i_j} 's is a characteristic map for the top-dimensional cell of P^I . We will be especially interested in the sequences $I = (i_1, \dots, i_m)$ satisfying $n > i_1 > \dots > i_m > 0$. These sequences will be called *admissible*, as will the sequence (0).

Proposition 6.5.1 ([22, Prop. 3D.1]). *The maps $\rho \circ \varphi^I : \mathbb{D}^I \rightarrow \mathrm{SO}_n$, for I ranging over all admissible sequences, are the characteristic maps of a CW-structure on SO_n for which the map $\rho : P^{n-1} \times P^{n-2} \times \dots \times P^1 \rightarrow \mathrm{SO}_n$ is cellular.*

In particular, there is a single 0-cell $e^0 = \{e\}$, so SO_n is path-connected. The other cells $e^I = e^{i_1} \times e^{i_2} \times \dots \times e^{i_m}$ are products, via the group operation in SO_n , of the cells $e^i \subset P^{n-1} \subset \mathrm{SO}_n$.

Proof. There are three properties we need to show in order to obtain a CW-structure:

- (1) For each decreasing sequence I , $\rho \circ \varphi^I$ is a homeomorphism from the interior of \mathbb{D}^I onto its image.
- (2) The resulting image cells e^I are all disjoint and cover SO_n .
- (3) For each e^I , $\rho \circ \varphi^I(\partial\mathbb{D}^I)$ is contained in a union of cells of lower dimension than e^I .

To begin the verification, define $p : \mathrm{SO}_n \rightarrow \mathbb{S}^{n-1}$ be evaluation at the vector $e_n = (0, \dots, 0, 1)$, $p(g) = g \cdot e_n$. Isometries in $P^{n-2} \subset P^{n-1} \subset \mathrm{SO}_n$ fix e_n , so $p(P^{n-2}) = \{e_n\}$. We claim that p is a homeomorphism from $P^{n-1} \setminus P^{n-2}$ onto $\mathbb{S}^{n-1} \setminus \{e_n\}$. This can be seen as follows. Thinking of a point in P^{n-1} as a vector v , the map p takes this to $p(v)(e_n) = r(v)r(e_1) \cdot e_n = r(v) \cdot e_n$ since $e_n \perp e_1$. From Figure 6.1 it is clear that p simply stretches the lower half of each meridian circle in \mathbb{S}^{n-1} onto the whole meridian circle, doubling the angle up from the south pole, so $P^{n-1} \setminus P^{n-2}$, represented by vectors whose last coordinate is negative, is taken homeomorphically into $\mathbb{S}^{n-1} \setminus \{e_n\}$.

The next statement is that the map of pairs

$$H : \left(P^{n-1} \times \mathrm{SO}_{n-1}, P^{n-2} \times \mathrm{SO}_{n-1} \right) \rightarrow (\mathrm{SO}_n, \mathrm{SO}_{n-1}), \quad h(v, g) = \rho(v)g, \quad (6.5.2)$$

is a homeomorphism from $(P^{n-1} \setminus P^{n-2}) \times \mathrm{SO}_{n-1}$ onto $\mathrm{SO}_n \setminus \mathrm{SO}_{n-1}$. Here we view SO_{n-1} as the subgroup of SO_n fixing the vector e_n . To construct an inverse to this homeomorphism, let $h \in \mathrm{SO}_n \setminus \mathrm{SO}_{n-1}$ be given. Then $h \cdot e_n \neq e_n$ so by the preceding paragraph there is a unique $v_h \in P^{n-1} \setminus P^{n-2}$ with $\rho(v_h) \cdot e_n = h \cdot e_n$, and v_h depends continuously on h since $h \cdot e_n$ does. The composition $g_h = \rho(v_h)^{-1}h$ then fixes h , hence lies in SO_{n-1} . Since $\rho(v_h)g_h = h$, the map $h \mapsto (v_h, g_h)$ is an inverse to H on $\mathrm{SO}_n \setminus \mathrm{SO}_{n-1}$.

Statements (1) and (2) can now be proven by induction on n . The map ρ takes P^{n-2} to SO_{n-1} , so we may assume inductively that the maps $\rho \circ \varphi^I$ for I ranging over admissible sequences with first term $i_1 < n - 1$ are the characteristic maps for a CW-structure on SO_{n-1} , with cells

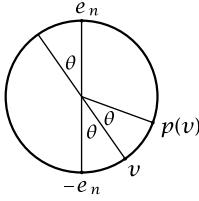


Fig. 6.1: Reflections defined in the proof of Proposition 6.5.1.

the corresponding products e^I . The admissible sequences I with $i_1 = n - 1$ then give disjoint cells e^I covering $\mathrm{SO}_n \setminus \mathrm{SO}_{n-1}$ by what was shown in the previous paragraph. So (1) and (2) hold for SO_n .

To prove (3) it suffices to show that there is an inclusion $P^i P^i \subset P^i P^{i-1}$ in SO_n since for an admissible sequence I , the map $\rho : P^I \rightarrow \mathrm{SO}_n$ takes the boundary of the top-dimensional cell of P^I to the image of products P^J with J obtained from I by decreasing one term i_j by 1, yielding a sequence which is admissible except perhaps for having two successive terms equal. As a preliminary to showing that $P^i P^i \subset P^i P^{i-1}$, observe that for $g \in O_n$ we have $r(g \cdot v) = gr(v)g^{-1}$. Hence $\rho(v)\rho(w) = r(v)r(e_1)r(w)r(e_1) = r(v)r(w')$ where $w' = r(e_1) \cdot w$. Thus to show $P^i P^i \subset P^i P^{i-1}$ it suffices to find for each pair $v, w \in \mathbb{R}^{i+1}$ a pair $x \in \mathbb{R}^{i+1}, y \in \mathbb{R}^i$ with $r(v)r(w) = r(x)r(y)$.

Let $V \subset \mathbb{R}^{i+1}$ be a 2-dimensional subspace containing v and w . Since $V \cap \mathbb{R}^i$ is at least 1-dimensional, we can choose a unit vector $y \in V \cap \mathbb{R}^i$. Let $g \in O_{i+1}$ take V to \mathbb{R}^2 and y to e_1 . Then the conjugate $gr(v)r(w)g^{-1} = r(g \cdot v)r(g \cdot w)$ lies in SO_2 , hence has the form $\rho(z) = r(z)r(e_1)$ for some $z \in \mathbb{R}^2$ by statement (2) for $n = 2$. Therefore

$$r(v)r(w) = g^{-1}r(z)r(e_1)g = r(g^{-1} \cdot z)r(g^{-1} \cdot e_1) = r(x)r(y) \quad (6.5.3)$$

for $x = g^{-1} \cdot z \in \mathbb{R}^{i+1}$ and $y \in \mathbb{R}^i$.

It remains to show that the map $\rho : P^{n-1} \times P^{n-2} \times \dots \times P^1 \rightarrow \mathrm{SO}_n$ is cellular. This follows from the inclusions $P^i P^i \subset P^i P^{i-1}$ derived above, together with another family of inclusions $P^i P^j \subset P^j P^i$ for $i < j$. To prove the latter we have the formulas

$$\begin{aligned} \rho(v)\rho(w) &= r(v)r(w') \text{ where } w' = r(e_1) \cdot w \\ &= r(v)r(w')r(v)r(v) \\ &= r(r(v) \cdot w')r(v) \text{ from } r(g \cdot v) = gr(v)g^{-1} \\ &= r(r(v)r(e_1) \cdot w)r(v) = r(\rho(v) \cdot w)r(v) \\ &= \rho(\rho(v) \cdot w)\rho(v') \text{ where } v' = r(e_1) \cdot v, \text{ hence } v = r(e_1) \cdot v'. \end{aligned} \quad (6.5.4)$$

In particular, taking $v \in \mathbb{R}^{i+1}$ and $w \in \mathbb{R}^{j+1}$ with $i < j$, we have $\rho(v) \cdot w \in \mathbb{R}^{j+1}$, and the product $\rho(v)\rho(w) \in P^i P^j$ equals the product $\rho(\rho(v) \cdot w)\rho(v') \in P^j P^i$. ■

This result can be used to easily compute the homology groups of SO_n , as is done in [22]. CW -structures also exist on SU_n and Sp_n . We don't describe them since for us it will only be important to know that they exist. Instead of using these CW -structures directly, we will use exact sequences of homotopy groups to compute *some* homotopy groups of the classical groups.

6.6 Lie group actions I

Definition 6.6.1 (Lie group action). Given a Lie group G and a smooth manifold M , a Lie group action of G on M is a smooth map $\Phi : G \times M \rightarrow M$ such that, denoting $\Phi_g = \Phi(g, _) : M \rightarrow M$, it holds that

$$\Phi_e = \mathrm{id}_M, \quad (6.6.1)$$

and either

$$\Phi_{gh} = \Phi_g \circ \Phi_h \quad (\text{left action}) \quad (6.6.2)$$

or

$$\Phi_{gh} = \Phi_h \circ \Phi_g \quad (\text{right action}). \quad (6.6.3)$$

In particular, each Φ_g is a diffeomorphism on M . As before, we will sometimes simply write $g \cdot x = \Phi(g, x)$ for left actions and $x \cdot g = \Phi(g, x)$ for right actions. The triple (M, G, Φ) is referred to as a Lie group action and the pair (M, Φ) as a G -manifold. When the name of the action is not needed, we will simply write $G \curvearrowright M$ for a left action and $M \curvearrowright G$ for a right one.

The notions of orbits, stabilizers, transitive actions and free actions transfer with no changes to the context of Lie groups.

Example 6.6.2 (Actions of G on itself). The left and right translation operators L_g, R_g on a Lie group G are smooth diffeomorphisms since $(L_g)^{-1} = L_{g^{-1}}$ is a smooth inverse. In addition,

$$L_{gh} = L_g \circ L_h, \quad R_{gh} = R_h \circ R_g. \quad (6.6.4)$$

Therefore L_g defines a left action of G on itself, and R_g defines a right action. Note that these two actions commute:

$$L_g \circ R_h = R_h \circ L_g. \quad (6.6.5)$$

In addition, the adjoint operator Ad_g defines the *adjoint action* of G on itself since

$$\text{Ad}_{gh} = \text{Ad}_g \circ \text{Ad}_h, \quad (6.6.6)$$

which makes it a left action. Similarly, $\text{Ad}_{g^{-1}}$ is a right action.

Example 6.6.3 (Lie group actions). (a) The natural action of $\text{GL}_n(V)$ on V via $(A, x) \mapsto Ax$ is a smooth left action.

- (b) An action of a discrete group Γ on a manifold is smooth iff every Φ_g is smooth. E.g. \mathbb{Z}^n acts smoothly and freely on \mathbb{R}^n by translations $(n, x) \mapsto x + n$.
- (c) The automorphism group $\text{Aut}(\pi)$ of a smooth covering map $\pi : E \rightarrow M$ is a zero-dimensional Lie group acting smoothly and freely on E .

Definition 6.6.4 (Equivariant maps). Let G be a Lie group, and let (M, Φ) and (M', Φ') be smooth manifolds endowed with smooth left or right G -actions Φ and Φ' . A map $F : M \rightarrow M'$ is called G -equivariant with respect to the given G -actions if

$$F \circ \Phi_g = \Phi'_g \circ F \quad (6.6.7)$$

for all $g \in G$. In other words, the following diagram must commute:

$$\begin{array}{ccc} M & \xrightarrow{F} & M' \\ \Phi_g \downarrow & & \downarrow \Phi'_g \\ M & \xrightarrow[F]{} & M'. \end{array} \quad (6.6.8)$$

This condition is also often expressed by saying F *intertwines* Φ and Φ' . Equivariant maps comprise the morphisms of the category $G\text{-Man}^\infty$ of smooth manifolds with a G -action. Furthermore, for two Lie group actions (M_i, G_i, Φ^i) , $i = 1, 2$, a morphism between them is a pair (φ, ρ) consisting of a smooth map $F : M_1 \rightarrow M_2$ and a Lie group homomorphism $\rho : G_1 \rightarrow G_2$ such that the corresponding square commutes:

$$F \circ \Phi^1 = \Phi^2 \circ (\rho \times F). \quad (6.6.9)$$

In this case we say that F intertwines the actions Φ^i , $i = 1, 2$.

Example 6.6.5. The map $\exp : \mathbb{R}^n \rightarrow \mathbb{T}^n$ given by $\exp(x^1, \dots, x^n) = (e^{2\pi i x^1}, \dots, e^{2\pi i x^n})$ is a smooth map that intertwines the actions of \mathbb{R} on \mathbb{R}^n and \mathbb{T}^n by

$$t \in \mathbb{R} \text{ acting on } \mathbb{R}^n : \quad t \cdot x = x + tv, \quad (6.6.10)$$

$$t \in \mathbb{R} \text{ acting on } \mathbb{T}^n : \quad (t \cdot z)^k = e^{2\pi i t v^k} z^k, \quad (6.6.11)$$

where $v \in \mathbb{R}^n$ is a fixed vector.

Theorem 6.6.6 (Equivariant rank theorem [7, Thm. 7.25]). *Let $F : (M, \Phi) \rightarrow (M', \Phi')$ be a smooth G -equivariant map and let the action Φ be transitive. Then F has constant rank.*

Proof. Let $p, q \in M$, and let $g \in G$ be such that $\Phi_g(p) = q$ (it exists by transitivity). By computing the differential of the equivariance condition, we get

$$T_q F \circ T_p \Phi_g = T_{F(p)} \Phi'_g \circ T_p F. \quad (6.6.12)$$

Since the differentials of group actions are isomorphisms of vector spaces, the other two maps have the same rank. This means that F has the same rank at any two points. ■

Definition 6.6.7 (Orbit map). Given a Lie group action $\Phi : G \times M \rightarrow M$, for each $p \in M$ we define the smooth orbit map

$$\Phi^p : G \rightarrow M, \quad g \mapsto \Phi(g, p). \quad (6.6.13)$$

Proposition 6.6.8 ([7, Prop. 7.26]). *The orbit map $\Phi^p : G \rightarrow M$ defined above has constant rank, so the stabilizer $G_p = (\Phi^p)^{-1}(p)$ is a properly embedded Lie subgroup of G . If $G_p = \{e\}$, then Φ^p is an injective smooth immersion, so the orbit $\Phi(G, p)$ is an immersed submanifold of M .*

Proof. First we note that Φ^p intertwines the left action of L_g on G and Φ on M . Since G acts transitively on itself, Φ^p has constant rank by Theorem 6.6.6. Thus G_p is properly embedded by Theorem 3.9.7 and a Lie subgroup by Proposition 6.3.2.

If $G_p = \{e\}$, then Φ^p is injective, and therefore by the Equivariant Rank Theorem 6.6.6 it is a smooth immersion (recall that an injective map of constant rank is an immersion by the Global Rank Theorem 3.7.5). This makes the orbit an immersed submanifold by Proposition 3.9.3. ■

Remark 6.6.9. We will later prove that the condition on the stabilizer subgroup can be dropped.

6.7 Semidirect products

Definition 6.7.1 (Semidirect product). A smooth left action $\Phi : H \times N \rightarrow N$ of a Lie group H on another Lie group N is called an action by automorphisms if the maps $\Phi_h : N \rightarrow N$ are group automorphisms for all $h \in H$.

Given such an action, the semidirect product of H and N is, denoted $N \rtimes_\Phi H$ as a manifold,

the Cartesian product $N \times H$, endowed with the multiplication

$$(n, h)(n', h') = (n\Phi_h(n'), hh'). \quad (6.7.1)$$

It is easy to check that the product defined above is a valid group product, and it's clearly smooth along with its inverse $(n, h)^{-1} = (\Phi_h^{-1}(n^{-1}), h^{-1})$.

Example 6.7.2 (The Euclidean group). $O_n(\mathbb{R})$ acts by automorphisms on the translation group \mathbb{R}^n . Their semidirect product is the Euclidean group $E_n = \mathbb{R}^n \rtimes O_n(\mathbb{R})$:

$$(b, A)(b', A') = (b + An', AA'). \quad (6.7.2)$$

It also acts on \mathbb{R}^n via

$$(b, A) \cdot x = b + Ax. \quad (6.7.3)$$

The following two theorems are a simple exercise.

Proposition 6.7.3 (Properties of semidirect products). *Let $G = N \rtimes_\Phi H$ as above.*

- (a) *The subsets $\tilde{N} = N \times \{e\}$ and $\tilde{H} = \{e\} \times H$ are closed Lie subgroups of G isomorphic to N and H , respectively.*
- (b) *\tilde{N} is a normal subgroup of G .*
- (c) *$\tilde{N} \cap \tilde{H} = \{(e, e)\}$ and $\tilde{N}\tilde{H} = G$.*

Theorem 6.7.4 (Characterization of semidirect products). *Let $N, H \subset G$ be closed Lie subgroups of a Lie group G such that N is normal, $N \cap H = \{e\}$, and $NH = G$. Then the map $(n, h) \mapsto nh$ is a Lie group isomorphism between $N \rtimes_{\text{Ad}} H$ and G , where $\text{Ad} : H \times N \rightarrow N$ is the adjoint action $\text{Ad}_h(n) = hn h^{-1}$.*

This kind of a semiproduct is called *internal*, but this theorem essentially shows that every semidirect product is internal with respect to the resulting group.

Exercise 6.6. Show the following isomorphisms:

- (a) $O_n(\mathbb{R}) \cong SO_n(\mathbb{R}) \rtimes O_1$.
- (b) $U_n(\mathbb{R}) \cong SU_n(\mathbb{R}) \rtimes U_1$.
- (c) $GL_n(\mathbb{R}) \cong SL_n(\mathbb{R}) \rtimes \mathbb{R}^\times$.
- (d) $GL_n(\mathbb{C}) \cong SL_n(\mathbb{C}) \rtimes \mathbb{C}^*$.

Prove that none of these groups are isomorphic to the analogous direct product groups by comparing the centers (the center is the set of all elements that commute with all of G).

6.8 Representations

Definition 6.8.1 (Linear representation). Let \mathbb{K} be either \mathbb{R} or \mathbb{C} . A \mathbb{K} -linear representation of a Lie group G is a \mathbb{K} -vector space V with a smooth left G -action Φ such that each map $\Phi_g : V \rightarrow V$ is a \mathbb{K} -linear automorphism. Equivalently, a representation is a smooth homomorphism $\rho = \Phi_\bullet : G \rightarrow \mathrm{GL}_\mathbb{K}(V)$.

A representation $\rho : G \rightarrow \mathrm{GL}(V)$ is called *faithful* if it is injective.

Note that in principle it is possible to define infinite-dimensional representations even of finite-dimensional groups, but it will take some extra work to introduce the proper topology and smooth structure in that case.

Example 6.8.2 (Lie group representations). (a) If G is any Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$ (such groups are called *matrix groups*), the inclusion map $G \hookrightarrow \mathrm{GL}_n(\mathbb{R})$ is a faithful representation, called the *defining representation* of G .

- (b) The inclusion map $\mathbb{S}^1 \hookrightarrow \mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$ is a faithful representation of the circle group.
- (c) Let $\sigma : \mathbb{R}^n \rightarrow \mathrm{GL}_{n+1}(\mathbb{R})$ be the map that sends $x \in \mathbb{R}^n$ to the matrix $\sigma(x)$ defined in block form by

$$\sigma(x) = \begin{pmatrix} I_n & x \\ 0 & 1 \end{pmatrix}, \quad (6.8.1)$$

where I_n is the $n \times n$ identity matrix and x is regarded as an $n \times 1$ column matrix. Then σ is a faithful representation of the additive group \mathbb{R}^n .

- (d) A faithful representation of the Euclidean group E_n is given by $\rho : E_n \rightarrow \mathrm{GL}_{n+1}(\mathbb{R})$ defined in block form by

$$\rho(b, A) = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}, \quad (6.8.2)$$

where b is considered as a column.

- (e) For positive integers n and d , let P_d^n denote the vector space of real polynomials $p : \mathbb{R}^n \rightarrow \mathbb{R}$ of degree at most d . For a matrix $A \in \mathrm{GL}_n(\mathbb{R})$ define the linear map $\tau(A) : P_d^n \rightarrow P_d^n$ by

$$\tau(A)p = p \circ A^{-1}. \quad (6.8.3)$$

This is a faithful representation of $\mathrm{GL}_n(\mathbb{R})$.

- (f) Another very important type of representation that is not linear is a projective representation, which is a homomorphism into the group $\mathrm{PGL}(V) = \mathrm{GL}(V)/(\mathbb{K}I)$ (where $\mathbb{K}I$ is the subgroup of multiplications by a scalar) of projective linear transformations of a vector space. The projective special linear group $\mathrm{PSL}_2(\mathbb{C}) = \mathrm{SL}_2(\mathbb{C})/\{I, -I\}$, also isomorphic to the proper orthochronous Lorentz group $\mathrm{SO}^+(1, 3)$, acts on the projective

complex line $\mathbb{C}P^1 \cong \mathbb{C} \cup \{\infty\}$ via Möbius transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}. \quad (6.8.4)$$

It is easy to check that this is a left action. It is *sharply 3-transitive*, meaning that for any triple of points (z, w, x) there exists a unique matrix $g \in \mathrm{PSL}_2(\mathbb{C})$ that maps them to $(0, 1, \infty)$.

Remark 6.8.3. Classical Lie groups are defined as groups of invertible matrices, and those matrices comprise the so called *defining representations* of those groups. Moreover, very often in introductory courses on Lie groups it is assumed that all groups under consideration are matrix groups. It is therefore natural to ask whether every Lie group is isomorphic to some matrix group (such groups are also called *linear*). The answer is no: not every finite-dimensional Lie group has a finite-dimensional faithful representation. The simplest example of a *nonlinear* group is $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, the universal covering group of $\mathrm{SL}_2(\mathbb{R})$. We will be able to show that it can't have a faithful representation later. Spin groups in even dimensions are also nonlinear. The fact of existence of such Lie groups is the ultimate justification for using general differential geometry to study them.

Every Lie group has a canonical representation on its own tangent space at the identity.

Definition 6.8.4 (Adjoint representation). The derivative of the adjoint action at the identity, $T_e \mathbf{Ad}_g : T_e G \rightarrow T_e G$, is a linear representation of G . We use the same symbol without boldface to denote it:

$$\mathrm{Ad}_g X = (\mathbf{Ad}_g)_{*e}(X). \quad (6.8.5)$$

To further differentiate them, we will omit parentheses when writing the linear map $\mathrm{Ad}_g X$ but keep them when writing $\mathbf{Ad}_g(h)$.

Note that any abelian Lie group has a trivial adjoint representation ($\mathrm{Ad}_g = 0$ for all $g \in G$).

Example 6.8.5 (Adjoint representation of matrix groups). Let $G \subset \mathrm{GL}(V)$ be a matrix group. Then $T_e G \cong \mathrm{Mat}(V)$ and we can explicitly compute the adjoint representation. Since $\mathrm{GL}(V)$ is open in $\mathrm{Mat}(V)$, we can define a straight path $\gamma(t) = A + tX$ for any $A \in \mathrm{GL}(V)$ and $X \in \mathrm{Mat}(V)$. In particular, let $h(t) = I + tX$. Then for any $A \in \mathrm{GL}(V)$, using formula (3.6.3) for the action of a differential of a smooth map,

$$\mathrm{Ad}_A X = \left. \frac{d}{dt} \right|_{t=0} A(I + tX)A^{-1} = AXA^{-1}. \quad (6.8.6)$$

In other words, on matrix groups the Lie algebraic version of the adjoint map is still just matrix conjugation.

6.9 Lie algebras: Construction I

Since Lie groups are manifolds, the next most natural thing to do is to study its tangent spaces and to see what kind of structure on them is induced by the group structure. First we note that since $L_g : G \rightarrow G$ is a diffeomorphism for every $g \in G$, it induces the isomorphisms

$$(L_g)_{*e} : T_e G \rightarrow T_g G \quad (6.9.1)$$

between the tangent spaces at the identity and any other point. Moreover, since $L_{gh} = L_g \circ L_h$, the chain rule for derivatives implies that $(L_g)_*$ is natural in the sense that it truly depends only on g itself and not how it might be decomposed into a product of other group elements. Since Lie group homomorphisms preserve the identity, their derivatives map $T_e G$ to $T_e H$. Therefore it will completely suffice to study the properties of the tangent space at the identity. First let us try to gain some intuition for how the group structure of G might be “encoded” in $T_e G$ by working in local coordinates. Suppose (U, φ) is a symmetric (i.e. $U^{-1} = U$) open neighborhood of $e \in G$ and $\varphi : U \rightarrow \mathbb{R}^n$ a chart map on it such that $\varphi(e) = 0$. We can find a smaller open neighborhood $U_0 \subset U$ such that the multiplication map $m : G \rightarrow G \rightarrow G$ restricts to $m : U_0 \times U_0 \rightarrow U$. The inversion map also restricts to $i : U \rightarrow U$. Then let $\mu : \varphi(U_0)^2 \rightarrow \varphi(U)$ and $\iota : \varphi(U) \rightarrow \varphi(U)$ be their representatives in the local chart. Since these are now smooth maps on Euclidean spaces, we can study their Taylor series at the origin. Since $m(e, g) = m(g, e) = g$ for all g , we have

$$\mu(0, x) = \mu(x, 0) = x \quad (6.9.2)$$

for all $x \in U_0$. This implies that both partial derivatives of μ computed at $(0, 0)$ are the identity. Therefore the Taylor series of μ around the origin has the form

$$\mu(x, y) = x + y + \beta(x, y) + \gamma(x, y) + O(4), \quad (6.9.3)$$

where $\beta(x, y) : \varphi(U_0)^2 \rightarrow \mathbb{R}^n$ is some vector-valued bilinear form of $x, y \in \mathbb{R}^n$, $\gamma(x, y)$ is a similar cubic term, and $O(4)$ stands for terms that are quartic in x and y together (x^4, x^3y , etc.). In addition, since $m(g, i(g)) = e$, we also have

$$\mu(x, \iota(x)) = \mu(\iota(x), x) = 0, \quad (6.9.4)$$

which implies that the Taylor series of the inversion map has the form

$$\iota(x) = -x + \beta(x, x) + O(x^3). \quad (6.9.5)$$

The extra constraints on the forms β and γ will come from the fact that they’re not just some vector-valued forms, but that due to the group structure $m : G \times G \rightarrow G$ the domain and codomain of these forms are closely connected to each other.

Lemma 6.9.1. *The antisymmetric quadratic form*

$$b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad b(x, y) = \beta(x, y) - \beta(y, x), \quad (6.9.6)$$

is natural with respect to local coordinate transformations on G . That is, for any matrix $A \in \mathrm{GL}_n(\mathbb{R})$,

$$Ab(x, y) = b(Ax, Ay). \quad (6.9.7)$$

Proof. Any local coordinate transformation on U_0 , represented by a transition diffeomorphism f with $f(0) = 0$, will transform the local multiplication map μ into $\tilde{\mu}$ with

$$f(\tilde{\mu}(x, y)) = \mu(f(x), f(y)). \quad (6.9.8)$$

Writing the Taylor series of f as

$$f(x) = Ax + f_2(x) + O(3), \quad (6.9.9)$$

where A is the Jacobi matrix of f at the origin and $f_2(x)$ is the quadratic term, we have

$$\mu(f(x), f(y)) = Ax + Ay + f_2(x) + f_2(y) + \beta(Ax, Ay) + O(3) \quad (6.9.10)$$

and

$$f(\tilde{\mu}(x, y)) = Ax + Ay + A\tilde{\beta}(x, y) + f_2(x + y) + O(3). \quad (6.9.11)$$

By comparing these two expressions we conclude that

$$f_2(x) + f_2(y) + \beta(Ax, Ay) = A\tilde{\beta}(x, y) + f_2(x + y). \quad (6.9.12)$$

In other words, the quadratic part of the multiplication map also depends on the quadratic part of the coordinate map itself. However, by noticing that the terms depending on f_2 are symmetric in x and y , we get the statement of the theorem. ■

This indicates the existence of a natural antisymmetric bilinear form $T_e G \times T_e G \rightarrow T_e G$. So far we have only used the property of the identity $eg = ge = g$. Aside from this, group multiplication also satisfies associativity.

Lemma 6.9.2. *The associativity of group multiplication implies that*

$$\beta(\beta(x, y), z) + \gamma(x, y) + \gamma(x + y, z) = \beta(x, \beta(y, z)) + \gamma(y, z) + \gamma(x, y + z). \quad (6.9.13)$$

Proof. We have

$$\begin{aligned}\mu(\mu(x, y), z) &= x + y + \beta(x, y) + \gamma(x, y) + z + \beta(x + y, z) + \beta(\beta(x, y), z) + \\ &\quad + \gamma(x + y, z) + O(4)\end{aligned}\quad (6.9.14)$$

and

$$\begin{aligned}\mu(x, \mu(y, z)) &= x + y + z + \beta(x, y + z) + \beta(y, z) + \gamma(y, z) + \beta(x, \beta(y, z)) + \\ &\quad + \gamma(x, y + z) + O(4).\end{aligned}\quad (6.9.15)$$

These two expressions must be equal by associativity. All terms of degree less than 3 cancel out because β is bilinear, and we get the claimed formula. ■

Lemma 6.9.3. *The antisymmetric form $b(x, y)$ satisfies the Jacobi identity*

$$b(x, b(y, z)) + b(y, b(z, x)) + b(z, b(x, y)) = 0. \quad (6.9.16)$$

Proof. We antisymmetrize the equation from Lemma 6.9.2 with respect to permutations of all three arguments. First we notice that the antisymmetrization of $\gamma(x, y)$ will be the same as that of $\gamma(y, z)$ since they are related to each other by a cyclic (and thus even) permutation. Moreover, the antisymmetrizations of both $\gamma(x + y, z)$ and $\gamma(x, y + z)$ vanish since both of these are symmetric under a transposition of two of the arguments, and the only function that is both symmetric and antisymmetric is zero.

What remains is

$$\mathcal{A}\beta(\beta(x, y), z) = \mathcal{A}\beta(x, \beta(y, z)), \quad (6.9.17)$$

where \mathcal{A} is the antisymmetrization operator

$$\mathcal{A}f(x_1, x_2, x_3) = \sum_{\sigma \in S_3} \text{sign}(\sigma) f(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}). \quad (6.9.18)$$

This equation has six terms on each side, and we are going to group them into three groups of four by gathering terms that have the same variable standing alone. We start with those having z as a lone variable. On the left hand side we combine two terms coming from the transposition $(1 \Leftrightarrow 2)$

$$\beta(\beta(x, y), z) - \beta(\beta(y, x), z) = \beta(b(x, y), z), \quad (6.9.19)$$

whereas on the right hand side we find

$$\beta(z, \beta(x, y)) - \beta(z, \beta(y, x)) = \beta(z, b(x, y)). \quad (6.9.20)$$

Finally, bringing these two expressions to the same side we find

$$\beta(b(x, y), z) - \beta(z, b(x, y)) = b(b(x, y), z). \quad (6.9.21)$$

The full antisymmetrization is just a sum of cyclic permutations of this one, which gives us the Jacobi identity. ■

Remark 6.9.4. It's always possible to find some local coordinates in which $\beta(x, y) = \frac{1}{2}b(x, y)$ is antisymmetric right away (namely the coordinates in which $\iota(x) = -x + O(x^3)$). Later we will prove the miraculous result that there is a special coordinate system in which $\iota(x) \equiv -x$ and not only the quadratic term, but *all* higher terms in the Taylor series of $\mu(x, y)$ can be written in terms of *only* $b(x, y)$, and there is a *universal* formula giving this dependence (the Baker-Campbell-Hausdorff formula). In other words, the entire local structure of the Lie group is encoded in the bilinear form b .

6.10 Lie algebras: Construction II

The above calculations can be circumvented if we display a manifestly functorial object that gives an antisymmetric bilinear form on $T_e G$. To this end, let us take a look at the local representation of the conjugation map $\text{Ad}_g h = ghg^{-1}$:

$$c(x, y) = \mu(\mu(x, y), \iota(x)). \quad (6.10.1)$$

Lemma 6.10.1. *The Taylor series of the local conjugation map reads*

$$c(x, y) = y + b(x, y) + O(3). \quad (6.10.2)$$

Proof. Combining the Taylor series of μ and ι , we have

$$\begin{aligned} \mu(\mu(x, y), \iota(x)) &= x + y + \iota(x) + \beta(x + y, \iota(x)) + \beta(x, y) + O(3) = \\ &= x + y - x + \beta(x, x) + \beta(x, -x) + \beta(y, -x) + \beta(x, y) + O(3) = \\ &= y + \beta(x, y) - \beta(y, x) + O(3) = y + b(x, y) + O(3). \end{aligned} \quad (6.10.3)$$

Moreover, since $c(x, 0) = 0$ for all x in some neighborhood of the origin, we know that the error term $O(3)$ vanishes when $y = 0$ and therefore is $O(3) = yO(x^2, xy, y^2)$. This means that the double derivative with respect to x and y , which naturally acts on two vectors $X, Y \in \mathbb{R}^n$ reads

$$\left(\frac{\partial^2}{\partial x \partial y} c(x, y) \right) (X, Y) = b(X, Y) + O(x, y)(X, Y). \quad (6.10.4)$$

At $x = y = 0$ the error term vanishes and we get

$$\left(\frac{\partial^2}{\partial x \partial y} \Big|_{x=y=0} c(x, y) \right) (X, Y) = b(X, Y). \quad (6.10.5)$$

Obviously this double derivative is nothing but the local chart representation of the double derivative of $\text{Ad}_g(h) = ghg^{-1}$ viewed as a map $G \times G \rightarrow G$ with respect to both variables. Recall that we call its derivative at the identity $(\text{Ad}_g)_{*e} : T_e G \rightarrow T_e G$ by the symbol Ad_g . We can now think of this as a map $\text{Ad}_{\bullet} : G \rightarrow \text{GL}(T_e G)$. With this, we have

$$T_{g=e} T_{h=e} (\text{Ad}_g(h)) = T_e \text{Ad}_{\bullet}. \quad (6.10.6)$$

Note that the two derivatives commute because partial derivatives of smooth maps on \mathbb{R}^n commute. This is a linear operator $T_e \text{Ad}_{\bullet} : T_e G \rightarrow T_e \text{GL}(T_e G) \cong \text{End}(T_e G)$ from the tangent space $T_e G$ to the space of linear operators on it (this can also be thought of as a tensor on $T_e G$ of contravariant rank 1 and covariant rank 2).

Definition 6.10.2 (Adjoint operator, Lie bracket). The linear adjoint operator is

$$\text{ad} = T_e \text{Ad} : T_e G \rightarrow \text{End}(T_e G). \quad (6.10.7)$$

More explicitly, for every $X, Y \in T_e G$, the action of $\text{ad}_X : T_e G \rightarrow T_e G$ is defined by

$$\text{ad}_X Y = (T_e(\text{Ad}_{\bullet} Y))(X). \quad (6.10.8)$$

Equivalently, using the definition of tangent vectors as equivalence classes of curves, we have

$$\begin{aligned} \text{ad}_X Y &= \frac{d}{dt} \frac{d}{ds} \Big|_{t=s=0} g(t)h(s)g^{-1}(t) = \frac{d}{dt} \frac{d}{ds} \Big|_{t=s=0} g(t)h(s)g^{-1}(t)h^{-1}(s) = \\ &= \frac{d}{dt} \frac{d}{ds} \Big|_{t=s=0} [g(t), h(s)]. \end{aligned} \quad (6.10.9)$$

for any two curves $g(t)$ and $h(s)$ such that $X = g'(0)$ and $Y = h'(0)$. The fact that inserting the extra $h^{-1}(s)$ doesn't change the result is due to the fact that it can't contribute anything to the $O(ts)$ term in the Taylor series since the $O(t)$ term in ghg^{-1} is already zero (to first order ghg^{-1} in local coordinates reads $tX + sY - tX = sY$ as follows from our local functions μ and i). This identity can be treated as yet another construction of the Lie algebra, but we already know that it is equivalent to the others.

Note that in local coordinates the anti-symmetric form that we derived in the first construction is exactly the Lie bracket written in a local chart (U, φ) :

$$b(\varphi_{*e} X, \varphi_{*e} Y) = \varphi_{*e}(\text{ad}_X Y). \quad (6.10.10)$$

The anti-symmetry and the Jacobi identity for $\text{ad}_X Y$ follow from those for $b(X, Y)$.

Definition 6.10.3 (Lie algebra). A Lie algebra is a pair $(\mathfrak{g}, [\cdot, \cdot])$ where \mathfrak{g} is a finite-dimensional vector space over the field \mathbb{K} and $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ is a map called a Lie bracket that satisfies the following properties for all $X, Y, Z \in \mathfrak{g}$:

1. bilinearity: for all $a, b \in \mathbb{K}$,

$$[aX + bY, Z] = a[X, Z] + b[Y, Z], \quad [X, aY + bZ] = a[X, Y] + b[X, Z]; \quad (6.10.11)$$

2. alternating property:

$$[X, X] = 0; \quad (6.10.12)$$

3. Jacobi identity:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \quad (6.10.13)$$

Note that over fields of characteristic other than 2, like \mathbb{R} or \mathbb{C} the last two properties together imply skew-symmetry (anticommutativity) $[X, Y] = -[Y, X]$.

A Lie subalgebra of \mathfrak{g} is a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$, i.e. the restriction of the Lie bracket to \mathfrak{h} produces a Lie bracket on \mathfrak{h} and the inclusion map is a homomorphism of Lie algebras. In this case we write $\mathfrak{h} \leqslant \mathfrak{g}$.

Example 6.10.4. Any associative algebra (A, \cdot) (i.e. a vector space with a binary associative multiplication operation) can be endowed with a structure of a Lie group by setting the Lie bracket as the commutator

$$[X, Y] = X \cdot Y - Y \cdot X. \quad (6.10.14)$$

For example, the algebra $L(V) = \text{End}(V)$ of linear maps on a vector space V admits this structure. In fact this is a forgetful functor $\text{AssAlg} \rightarrow \text{LieAlg}$.

Historically, most Lie algebras were “discovered” as not just the Lie algebras associated with associative rings (i.e. commutator algebras), but as algebras of derivations on rings.

Definition 6.10.5 (Derivation). Given a (not necessarily associative) ring R , a derivation on it is a map $D : R \rightarrow R$ that is additive ($D(x + y) = Dx + Dy$) and satisfies the Leibniz rule $D(xy) = Dx \cdot y + x \cdot Dy$.

Example 6.10.6 (Lie ring/algebra of derivations). Denote the set of all derivations on a ring R by $\text{Der}(R)$. Clearly the sum of two derivations is a derivation, however the composition is not always a derivation. However, it is easy to verify that the commutator $[D_1, D_2] = D_1 D_2 - D_2 D_1$ of two derivations is a derivation. Thus $\text{Der}(R)$ has a natural structure of a Lie ring. If $R = A$ is actually an algebra over some field, then $\text{Der}(A)$ becomes a Lie algebra. As we will soon see, when A is an algebra of smooth scalar functions, $\text{Der}(A)$

consists of a certain type of differential operators.

Definition 6.10.7 (Lie algebra homomorphism). Given two Lie algebras $\mathfrak{g}, \mathfrak{g}'$, a homomorphism $\phi : \mathfrak{g} \rightarrow \mathfrak{g}'$ is a linear map such that

$$\phi[X, Y]_{\mathfrak{g}} = [\phi X, \phi Y]_{\mathfrak{g}'} . \quad (6.10.15)$$

These maps comprise the morphisms in the category $\text{LieAlg}_{\mathbb{K}}$ of Lie algebras over the field \mathbb{K} .

Example 6.10.8 (Inner derivations on a Lie algebra). On any abstract Lie algebra $(\mathfrak{g}, [\cdot, \cdot])$ one can define the linear adjoint operators $\text{ad}_x : \mathfrak{g} \rightarrow \mathfrak{g}$ by $\text{ad}_x(y) = [x, y]$. It is then easy to see that the Jacobi identity for $[\cdot, \cdot]$ is equivalent to the statement that ad_x is a derivation on \mathfrak{g} :

$$\text{ad}_x([y, z]) = [\text{ad}_x y, z] + [y, \text{ad}_x z]. \quad (6.10.16)$$

Moreover, since $\text{ad}_x y = -\text{ad}_y x$, this can also be written as

$$\text{ad}_{[x, y]} z = [\text{ad}_x, \text{ad}_y] z, \quad (6.10.17)$$

which is equivalent to saying that $\text{ad} : \mathfrak{g} \rightarrow \text{gl}(\mathfrak{g})$ is a *homomorphism* of Lie algebras. Therefore the set of all adjoint operators comprises a Lie subalgebra $\text{Inn}(\mathfrak{g}) \subset \text{gl}(\mathfrak{g})$, called the *algebra of inner derivations* on \mathfrak{g} .

Knowing from Construction I that $\text{ad}_X Y$ satisfies the properties of a Lie bracket, we can finally define the Lie algebra corresponding to a Lie group.

Definition 6.10.9 (Lie algebra of a Lie group). Given a Lie group G , its Lie algebra is $\mathfrak{g} = T_e G$ with the Lie bracket $[X, Y] = \text{ad}_X Y$.

The Jacobi identity can be alternatively written as

$$\text{ad}_X [Y, Z] = [\text{ad}_X Y, Z] + [Y, \text{ad}_X Z], \quad (6.10.18)$$

which is equivalent to saying that the adjoint ad_X is a derivation with respect to the Lie bracket.

Lemma 6.10.10. *If $G = \text{GL}(V)$ and $X, Y \in T_e G \cong \text{End}(V)$, then*

$$\text{ad}_X Y = [X, Y] = XY - YX. \quad (6.10.19)$$

Proof. As we've shown before, on matrix groups $\text{Ad}_A Y = AYA^{-1}$. Letting $A = I + tX$ and noting that $A^{-1} = I - tX + O(t^2)$, we find

$$\text{Ad}_{I+tX} Y = t(XY - YX) + O(t^2), \quad (6.10.20)$$

which upon differentiation at $t = 0$ implies the claimed formula. ■

Example 6.10.11 ($\mathfrak{gl}_n(R)$). The so called *general linear Lie algebra* constructed in Lemma 6.10.10 is denoted $\mathfrak{gl}(V) = \text{End}(V)^{(-)}$ (recall that $R \mapsto R^{(-)}$ is the forgetful functor that turns a ring into its associated Lie ring by defining the Lie bracket as the commutator). In fact it can be defined over all associative rings. Usually we use the standard basis consisting of the *standard matrix units* e_{ij} , $i \leq j$. Since $e_{jh}e_{ik} = \delta_{jh}e_{ik}$, the commutator reads

$$[e_{ij}, e_{hk}] = \delta_{jh}e_{ij} - \delta_{ki}e_{hk}. \quad (6.10.21)$$

Later we will see that the Poincaré-Birkhoff-Witt Theorem implies that any Lie algebra L is isomorphic to a subalgebra of some Lie algebra $A^{(-)}$ associated with some associative algebra A . Moreover, the theorem of Ado and Iwasawa implies that for a finite-dimensional L the algebra A can be taken to be finite-dimensional.

Example 6.10.12 ($\mathfrak{sl}_n(R)$). The commutator of any two traceless matrices is traceless, thus we have the Lie subalgebra of $\mathfrak{gl}_n(R)$ consisting of traceless matrices:

$$\mathfrak{sl}_n(R) = \{x \in \mathfrak{gl}_n(R) \mid \text{tr } x = 0\}, \quad (6.10.22)$$

called the *special linear Lie algebra*.

Exercise 6.7. Come up with a way to embed $\mathfrak{gl}_n(R)$ as a Lie subalgebra of $\mathfrak{sl}_{n+1}(R)$.

Now it is instructive to repeat the derivation of the properties of this Lie bracket without reference to local coordinates. First, we prove its naturality.

Proposition 6.10.13 ([33, Prop. 1.1.3]). *Let G, G' be Lie groups with corresponding Lie algebras $\mathfrak{g}, \mathfrak{g}'$, respectively, and suppose that $\Phi : G \rightarrow G'$ is a group homomorphism that is differentiable at $e \in G$. Write $\phi = T_e\Phi$. Then:*

1. $\phi \circ \text{Ad}_g = \text{Ad}_{\Phi(g)} \circ \phi$ for any $g \in G$.
2. $\phi : \mathfrak{g} \rightarrow \mathfrak{g}'$ is a homomorphism of Lie algebras.

Proof. Differentiating:

$$\Phi(\text{Ad}_g(h)) = \Phi(ghg^{-1}) = \Phi(g)\Phi(h)\Phi(g)^{-1} = \text{Ad}_{\Phi(g)}(\Phi(h)) \quad (6.10.23)$$

with respect to h at the identity in the direction of Y , we get

$$\Phi_{*e}\text{Ad}_g Y = \text{Ad}_{\Phi(g)}\Phi_{*e} Y. \quad (6.10.24)$$

Differentiating again with respect to g at the identity in the direction of X we get

$$\Phi_{*e}([\mathbf{X}, \mathbf{Y}]_{\mathfrak{g}}) = [\Phi_{*e}X, \Phi_{*e}Y]_{\mathfrak{g}'} \quad (6.10.25)$$

as claimed. ■

Corollary 6.10.14. *The above Proposition shows that we have constructed a functor called the Lie functor:*

$$\text{Lie} : \text{LieGr} \rightarrow \text{LieAlg} \quad (6.10.26)$$

that takes Lie groups to their Lie algebras and Lie group homomorphisms to their derivatives at the identity.

Remark 6.10.15. Following convention, we will use small Gothic letters to denote Lie algebras, for example $\mathfrak{g} = \text{Lie } G$.

Remark 6.10.16. Looking ahead, we will ultimately show that the Lie functor, under proper restrictions, becomes an equivalence of categories. Namely, homomorphisms between *connected* Lie groups are uniquely determined by their derivatives at e , and every finite-dimensional Lie algebra has a unique simply connected Lie group corresponding to it. Taken together, these statements were originally proven by Sophus Lie and are sometimes called Lie's Fundamental Theorems.

Now we are finally ready show (once again) that $\text{ad}_X Y$ indeed satisfies the properties of a Lie bracket.

Theorem 6.10.17. *Let G be a Lie group and let $[X, Y] = \text{ad}_X Y$ be the Lie bracket in the Lie algebra $\mathfrak{g} = \text{Lie } G$. Then for all $X, Y, Z \in \mathfrak{g}$:*

$$[X, Y] = -[Y, X], \quad (6.10.27)$$

$$[[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]]. \quad (6.10.28)$$

Proof. $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$ is a smooth homomorphism, hence Proposition 6.10.13 shows that $\text{ad} = T_e \text{Ad}$ is a homomorphism of Lie algebras $\mathfrak{g} \rightarrow L(\mathfrak{g})$ (as usual, $L(V)$ denotes the space of all linear maps $V \rightarrow V$ in the category of vector spaces). Recall now that the Lie algebra on $\text{End}(\mathfrak{g})$ is given by the commutator, therefore

$$\text{ad}_{[X, Y]} = \text{ad}_X \text{ad}_Y - \text{ad}_Y \text{ad}_X. \quad (6.10.29)$$

Applying this to $Z \in \mathfrak{g}$ we get the Jacobi identity. For skew-symmetry we observe that the group commutator

$$(g, h) \mapsto [g, h] = ghg^{-1}h^{-1} \quad (6.10.30)$$

defines a smooth mapping $G \times G \rightarrow G$ whose derivative at (e, e) is zero (this expresses the fact that $X + Y = Y + X$ in \mathfrak{g}). As a consequence, its second-order derivative at (e, e) is an intrinsically defined symmetric bilinear mapping from $(\mathfrak{g} \times \mathfrak{g}) \times (\mathfrak{g} \times \mathfrak{g})$ into \mathfrak{g} . The

derivative of $[g, h]$ with respect to h at e in the direction of $Y \in \mathfrak{g}$ is equal to $(\text{Ad}_g - I)(Y)$. Then the derivative of this with respect to g at e in the direction of $X \in \mathfrak{g}$ is equal to $\text{ad}_X Y$. Applying these differentiations in the opposite order, we first get $(I - \text{Ad}_h)(X)$, and then $-\text{ad}_Y X$. Since the order of these differentiations doesn't actually matter for C^2 maps, we conclude skew-symmetry. ■

Example 6.10.18 (Lie algebras of $\text{SO}_3(\mathbb{R})$ and SU_2). Continuing Example 6.1.10, the Lie algebra of $\text{SO}_3(\mathbb{R})$ can clearly be identified with the Lie algebra of antisymmetric 3×3 real matrices. Moreover, it can be identified with the Lie algebra (\mathbb{R}^3, \times) , where \times is the vector product, via the isomorphism $\omega \mapsto A_\omega$ (see formula (6.1.7)).

The Lie algebra of SU_2 , denoted \mathfrak{su}_2 , can be identified with the space

$$\mathfrak{su}_2 = \left\{ \begin{pmatrix} i\alpha & \beta \\ -\bar{\beta} & -i\alpha \end{pmatrix} \middle| \alpha \in \mathbb{R}, \beta \in \mathbb{C} \right\} \quad (6.10.31)$$

of anti-selfadjoint traceless complex 2×2 matrices. We identify this space with $\mathbb{R}^3 \cong \mathbb{R} \times \mathbb{C}$. Using the notation from above for an element $g(a, b) \in \text{SU}_2$, we get

$$\text{Ad}_{g(a,b)} = \begin{pmatrix} |a|^2 - |b|^2 & 2 \operatorname{Im}(a\bar{b}) & 2 \operatorname{Re}(a\bar{b}) \\ 2 \operatorname{Im}(ab) & \operatorname{Re}(a^2 + b^2) & -\operatorname{Im}(a^2 - b^2) \\ -2 \operatorname{Re}(ab) & \operatorname{Im}(a^2 + b^2) & \operatorname{Re}(a^2 - b^2) \end{pmatrix}. \quad (6.10.32)$$

This can be identified with a rotation matrix R_n with

$$n = \operatorname{sign}(\operatorname{Re} a) \frac{\arccos(2(\operatorname{Re} a)^2 - 1)}{\sqrt{1 - (\operatorname{Re} a)^2}} \begin{pmatrix} \operatorname{Im} a \\ \operatorname{Re} b \\ \operatorname{Im} b \end{pmatrix}. \quad (6.10.33)$$

It follows that the adjoint map defines a group homomorphism $\text{Ad} : \text{SU}_2 \rightarrow \text{SO}_3(\mathbb{R})$ with kernel equal to $\{\pm I\} \cong \mathbb{Z}_2$, therefore this is a two-fold covering of $\text{SO}_3(\mathbb{R})$ with the 3-sphere, where antipodal points of the sphere are mapped to the same element of $\text{SO}_3(\mathbb{R})$. This once again leads to the diffeomorphism $\text{SO}_3(\mathbb{R}) \cong \mathbb{RP}^3$.

Example 6.10.19. As an example of how functoriality simplifies computations, recall the adjoint representation $\text{Ad}_g = (\text{Ad}_g)_{*e} : \mathfrak{g} \rightarrow \mathfrak{g}$. Since Ad_g is a Lie group isomorphism, we automatically conclude that Ad_g is a Lie algebra isomorphism, i.e.:

$$\text{Ad}_g [X, Y] = [\text{Ad}_g X, \text{Ad}_g Y]. \quad (6.10.34)$$

Example 6.10.20 (Adjoint representation of $\text{SO}_3(\mathbb{R})$). As we saw in Example 6.10.18, the adjoint representation $\text{Ad} : \text{SU}_2 \rightarrow \text{GL}(\mathfrak{su}_2)$ can be seen as a homomorphism $\text{Ad} : \text{SU}_2 \rightarrow$

$\mathrm{SO}_3(\mathbb{R})$. Therefore its derivative $\mathrm{ad} : \mathfrak{su}_2 \rightarrow \mathfrak{gl}(\mathfrak{su}_2)$ can be seen as $\mathrm{ad} : \mathfrak{su}_2 \rightarrow \mathfrak{so}_3$. By linearizing formula (6.10.32) near the identity, i.e. by Taylor expanding it at a point

$$g = \begin{pmatrix} 1 + i\alpha & \beta \\ -\bar{\beta} & 1 - i\alpha \end{pmatrix} + O(\alpha^2, |\beta|^2), \quad (6.10.35)$$

we get

$$\mathrm{ad} : \begin{pmatrix} i\alpha & \beta \\ -\bar{\beta} & -i\alpha \end{pmatrix} \mapsto \begin{pmatrix} 0 & -2 \operatorname{Im} \beta & 2 \operatorname{Re} \beta \\ 2 \operatorname{Im} \beta & 0 & -2\alpha \\ -2 \operatorname{Re} \beta & 2\alpha & 0 \end{pmatrix}. \quad (6.10.36)$$

The factor of 2 here is a manifestation of the fact that SU_2 is a double cover of SO_3 .

6.11 Lie subalgebras and ideals

Assume now that $i : H \rightarrow G$ is an immersive Lie group homomorphism. By using translation isomorphisms this is equivalent to $i_{*e} : \mathrm{Lie} H \rightarrow \mathrm{Lie} G$ being injective. By functoriality it respects the Lie bracket, which means that i_{*e} identifies $\mathrm{Lie} H$ with a Lie subalgebra of $\mathrm{Lie} G$. In other words, the following Proposition holds.

Proposition 6.11.1. *If $H < G$ is an immersed Lie subgroup of a Lie group G , then $\mathrm{Lie} H$ is a Lie subalgebra of $\mathrm{Lie} G$.*

The obvious question is now: Does the converse hold? The answer turns out to be yes, but we will have to wait to prove it.

Example 6.11.2. $\mathrm{Lie}(G \times H) \cong \mathrm{Lie} G \oplus \mathrm{Lie} H$.

Definition 6.11.3 (Lie algebra representation). A representation of a Lie algebra \mathfrak{g} over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ is a \mathbb{K} -vector space V with a \mathbb{K} -linear homomorphism of Lie algebras $\rho : (\mathfrak{g}, [\cdot, \cdot]) \rightarrow \mathfrak{gl}(V)$ where the Lie bracket on $\mathfrak{gl}(V) \cong \mathrm{End}(V)$ is the canonical commutator bracket.

Definition 6.11.4 (Adjoint representation of a Lie algebra). For any Lie algebra \mathfrak{g} , its adjoint representation is defined by the homomorphism

$$\mathrm{ad} : \mathfrak{g} \rightarrow \mathrm{Der}(\mathfrak{g}) \leq \mathfrak{gl}(\mathfrak{g}), \quad X \mapsto \mathrm{ad}_X. \quad (6.11.1)$$

The image of this homomorphism is $\mathrm{Inn}(\mathfrak{g})$, the algebra of inner derivations.

Example 6.11.5 (The Heisenberg group and algebra). Let H_3 be the group of real 3×3

matrices of the form

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}. \quad (6.11.2)$$

Its Lie algebra has the basis X, Y, Z where each of these matrices is zero except for a 1 in the position of a, b, c , respectively. The Lie bracket on this basis is $[X, Y] = Z$, $[Y, Z] = [X, Z] = 0$.

Let $C^\infty(\mathbb{R})$ be the vector space of all smooth functions on \mathbb{R} . We define a representation of the Heisenberg algebra $\mathfrak{h}_3 = \text{Lie } H_3$ on this space via the following linear operators;

$$\rho(X)f = \frac{d}{dx}f, \quad \rho(Y)f = xf, \quad \rho(Z)f = f. \quad (6.11.3)$$

It is easy to check that this is indeed a representation of \mathfrak{h} . In quantum mechanics the Heisenberg algebra consists of the position and momentum operators plus the identity.

Exercise 6.8. The standard basis for $\mathfrak{sl}_2(\mathbb{K})$ is

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (6.11.4)$$

Compute the images of these vectors in the adjoint representation.

- Example 6.11.6.**
1. Every 1-dimensional subspace of a Lie algebra is a Lie subalgebra.
 2. All upper triangular matrices ($X_{ij} = 0$ for $i > j$) form a subalgebra $\mathfrak{b}_n(\mathbb{K})$ of $\mathfrak{gl}_n(\mathbb{K})$. We will later identify this is the Borel subalgebra of $\mathfrak{gl}_n(\mathbb{K})$.
 3. All strictly upper triangular matrices ($X_{ij} = 0$ for $i \geq j$) form a subalgebra $\mathfrak{n}_n(\mathbb{K})$ of $\mathfrak{gl}_n(\mathbb{K})$.
 4. All diagonal matrices form an abelian subalgebra $\mathfrak{d}_n(\mathbb{K})$ of $\mathfrak{gl}_n(\mathbb{K})$. We will later identify this is the Cartan subalgebra of $\mathfrak{gl}_n(\mathbb{K})$.

Definition 6.11.7 (Ideal). An ideal of a Lie algebra \mathfrak{g} is a linear subspace $\mathfrak{n} \subset \mathfrak{g}$ such that

$$[\mathfrak{g}, \mathfrak{n}] \subset \mathfrak{n}. \quad (6.11.5)$$

In other words, \mathfrak{n} must be stable under all inner derivations of \mathfrak{g} . Note that an ideal is automatically a Lie subalgebra, but in this case we use the special notation $\mathfrak{n} \trianglelefteq \mathfrak{g}$.

Proposition 6.11.8. *If $H \subset G$ is a normal immersed Lie subgroup, then $\text{Lie } H \trianglelefteq \text{Lie } G$ is an ideal.*

Proof. If H is normal, $\text{Ad} : G \rightarrow \text{GL}(T_e G)$ takes values in the subgroup of $\text{GL}(T_e G)$ that

leaves $T_e H$ invariant. Hence the same applies to its derivative ad. ■

Definition 6.11.9 (Normalizer of a Lie algebra). The normalizer of a subset $S \subset \mathfrak{g}$ is

$$N_{\mathfrak{g}}(S) = \{X \in \mathfrak{g} \mid [X, S] \subset S\}. \quad (6.11.6)$$

If S is a Lie subalgebra, then this is exactly the largest Lie subalgebra in which S is an ideal. A subalgebra is called self-normalizing if $N_{\mathfrak{g}}(S) = S$.

Exercise 6.9. 1. Prove that $\mathfrak{b}_n(\mathbb{K})$ and $\mathfrak{d}_n(\mathbb{K})$ are self-normalizing subalgebras of $\mathfrak{gl}_n(\mathbb{K})$. Calculate the normalizer of $\mathfrak{n}_n(\mathbb{K})$.
2. $\mathfrak{sl}_n(\mathbb{K})$ is an ideal of $\mathfrak{gl}_n(\mathbb{K})$.

Definition 6.11.10 (Center of a Lie algebra). The center of a Lie algebra \mathfrak{g} is the set

$$C(\mathfrak{g}) = \{X \in \mathfrak{g} \mid [X, \mathfrak{g}] = \{0\}\}. \quad (6.11.7)$$

It is automatically a Lie subalgebra of \mathfrak{g} by virtue of being the kernel of the Lie algebra homomorphism $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$.

Definition 6.11.11 (Centralizer). The centralizer of a subset $S \subset \mathfrak{g}$ is the set

$$C_{\mathfrak{g}}(S) = \{X \in \mathfrak{g} \mid [X, S] = \{0\}\}. \quad (6.11.8)$$

It is automatically a Lie subalgebra.

Exercise 6.10. 1. The centralizer of an ideal in a Lie algebra is also an ideal.
2. The kernel of any Lie algebra homomorphism is an ideal.
3. If $A, B \trianglelefteq \mathfrak{g}$ are two ideals, then $A + B$ and $A \cap B$ are also ideals.
4. If $A, B \trianglelefteq \mathfrak{g}$ are two ideals, then the commutator subspace $[A, B] = \{[a, b] \mid a \in A, b \in B\}$ is itself an ideal. This is in fact equivalent to the Jacobi identity.
5. For $A, B, C \trianglelefteq \mathfrak{g}$, verify

$$[A + B, C] = [A, C] + [B, C]. \quad (6.11.9)$$

Is it true that $[A \cap B, C] = [A, C] \cap [B, C]$?

6. For $A, B, C \trianglelefteq \mathfrak{g}$, verify

$$[[A, B], C] = [[B, C], A] + [[C, A], B]. \quad (6.11.10)$$

7. $\text{Inn}(\mathfrak{g})$ is an ideal of $\text{Der}(\mathfrak{g})$.
8. The normalizer of the Lie algebra of a Lie group G is the Lie algebra of the normalizer of G .

Definition 6.11.12 (Quotient algebra). For an ideal $\mathfrak{n} \trianglelefteq \mathfrak{g}$, one can define the quotient Lie algebra $\mathfrak{g}/\mathfrak{n}$ as the quotient of vector spaces (whose elements are all the subspaces of the form $X + \mathfrak{n}$) with the Lie bracket

$$[X + \mathfrak{n}, Y + \mathfrak{n}] = [X, Y] + \mathfrak{n}. \quad (6.11.11)$$

Definition 6.11.13 (Simple Lie algebra). A Lie algebra \mathfrak{g} is called simple if it has no proper ideals (that is, ideals other than the trivial ones, 0 and \mathfrak{g}).

Exercise 6.11. Show that the center $C(\mathfrak{g})$ cannot have codimension 1. What is the group-theoretic analog of this result?

Definition 6.11.14 (Semidirect product of Lie algebras). Let $\mathfrak{h} \leqslant \mathfrak{g}$ and $\mathfrak{n} \trianglelefteq \mathfrak{g}$. We say that \mathfrak{g} is the semidirect product of \mathfrak{h} and \mathfrak{n} if the restriction of the canonical projection $\mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{n}$ to \mathfrak{h} induces an isomorphism $\mathfrak{h} \cong \mathfrak{g}/\mathfrak{n}$.

Every semidirect product can be constructed as the modification of the direct sum of the Lie algebras $\mathfrak{h} \oplus \mathfrak{n}$ where for any $X \in \mathfrak{h}$ and $Y \in \mathfrak{n}$ we set $[X, Y] = \theta(X)Y$ with some Lie algebra homomorphism $\theta : \mathfrak{h} \rightarrow \text{Der}(\mathfrak{n})$.

Example 6.11.15. The algebra $\mathfrak{b}_n(\mathbb{K})$ of triangular matrices is the semi-direct product of the ideal $\mathfrak{n}_n(\mathbb{K})$ and the complementary subalgebra $\mathfrak{d}_n(\mathbb{K})$.

6.12 Low-dimensional Lie algebras

Recall that an ideal is a subspace that is stable under only inner derivations on \mathfrak{g} . Moreover, $A \trianglelefteq B \trianglelefteq C$ doesn't imply $A \trianglelefteq C$. These defects are corrected in the next concept.

Definition 6.12.1 (Characteristic ideal). A subspace $\mathfrak{n} \subset \mathfrak{g}$ is called a characteristic ideal of the Lie algebra \mathfrak{g} if

$$\text{Der}(\mathfrak{g}) \cdot \mathfrak{n} \subset \mathfrak{n}, \quad (6.12.1)$$

i.e. it is stable under all derivations of \mathfrak{g} . It is automatically an ideal.

Exercise 6.12.

1. A characteristic ideal of an ideal is an ideal.
2. A characteristic ideal of a characteristic ideal is a characteristic ideal.
3. If A, B are two characteristic ideals of \mathfrak{g} , then $[A, B]$ is one too.
4. If $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is such that $[\mathfrak{n}, \mathfrak{n}] = \mathfrak{n}$, then it is characteristic.

Definition 6.12.2 (Structure constants). Let e_1, \dots, e_n be a basis of an n -dimensional Lie algebra \mathfrak{g} over a field \mathbb{K} . Then

$$[e_i, e_j] = \sum_k c_{ij}^k e_k \quad (6.12.2)$$

for some set of numbers $c_{ij}^k \in \mathbb{K}$ called the structure constants of \mathfrak{g} .

Exercise 6.13. Write down the two constraints on the structure constants imposed by the structure of the Lie algebra.

Example 6.12.3 (2-dimensional Lie algebra). Up to isomorphism, there is clearly only one non-abelian Lie algebra of dimension 2. By scaling we can assume that it is spanned by two elements x, y such that $[x, y] = y$. Thus any such Lie algebra is isomorphic to the algebra of matrices

$$M = \left\{ \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \right\}, \quad (6.12.3)$$

with $x = e_{11}, y = e_{12}$. From the viewpoint of Lie theory, this is the Lie algebra of the $(ax + b)$ -group, the affine group of a line.

Lemma 6.12.4. All derivations of the two-dimensional non-abelian Lie algebra M are inner.

Proof. Let $D \in \text{Der}(M)$. Since $[M, M]$ is a characteristic ideal, $Dy \in [M, M]$ and thus $Dy = \lambda y$ for some $\lambda \in \mathbb{K}$. However, $\text{ad}_{\lambda x} y = \lambda y$ so that $(D - \text{ad}_{\lambda x})y = 0$. Thus, modulo inner derivations, we can from the very start assume that $Dy = 0$. Now

$$[Dx, y] = [Dx, y] + [x, Dy] = D[x, y] = Dy = 0 \quad (6.12.4)$$

and thus $Dx = \mu y$ for some $\mu \in \mathbb{K}$. But then $D = \text{ad}_{-\mu y}$ is itself an inner derivation. ■

Lemma 6.12.5. If the two-dimensional Lie algebra M is an ideal $M \trianglelefteq \mathfrak{g}$, then it is a direct summand of \mathfrak{g} .

Proof. As we know, $\mathfrak{n} = C_{\mathfrak{g}}(M)$ is an ideal of \mathfrak{g} . Since the center of M is trivial, $\mathfrak{n} \cap M = 0$. On the other hand, for any $x \in \mathfrak{g}$ the restriction $\text{ad}_x|_M$ is a derivation of M and thus, by the preceding lemma, is inner. Thus there exists $y \in M$ such that $\text{ad}_x|_M = \text{ad}_y|_M$. But then $\text{ad}_{x-y}(M) = 0$, hence $x - y \in \mathfrak{n}$ and we have the decomposition $x = (x - y) + y \in \mathfrak{n} \oplus M$. ■

Now we are ready to classify all 3-dimensional Lie algebras. Let $s = \dim[\mathfrak{g}, \mathfrak{g}]$.

- If $s = 0$, then \mathfrak{g} is abelian.
- If $s = 1$, we can assume that $[\mathfrak{g}, \mathfrak{g}]$ is generated by z . We distinguish two cases. If z is non-central, then \mathfrak{g} is the direct sum of the non-abelian 2-dimensional algebra M constructed above, and of the 1-dimensional algebra. If z is central, then exist x, y such that $[x, y] = z$. Clearly x, y, z must be linearly independent, so the algebra is isomorphic to the algebra of strictly upper-triangular matrices $\mathfrak{n}_3(\mathbb{K})$, which is the Heisenberg algebra $H_3(\mathbb{K})$.

- If $s = 2$, then $[\mathfrak{g}, \mathfrak{g}]$ must be abelian: if it isn't, by Lemma 6.12.5 it must be a direct summand, and thus $s = 1$, a contradiction. Thus we may assume that $[\mathfrak{g}, \mathfrak{g}]$ is generated by y, z with $[y, z] = 0$. Let $x \in \mathfrak{g} \setminus [\mathfrak{g}, \mathfrak{g}]$. Then

$$[x, y] = ay + bz, \quad [x, z] = cy + dz \quad (6.12.5)$$

for some $a, b, c, d \in \mathbb{K}$ such that $ad - bc \neq 0$. Clearly one can replace the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ by a conjugate (which corresponds to a change of basis on $[\mathfrak{g}, \mathfrak{g}]$) or scale it (corresponding to rescaling x). Thus, when \mathbb{K} is algebraically closed, this matrix can be reduced to one of the following Jordan forms:

$$\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \quad (6.12.6)$$

In other words, in the first case, there is a basis x, y, z such that

$$[x, y] = y, \quad [x, z] = \lambda z, \quad [y, z] = 0, \quad (6.12.7)$$

for some $\lambda \in \mathbb{K}^*$, and two such algebras with different λ 's are isomorphic iff $\lambda_1 = \lambda_2^{\pm 1}$. For an infinite field this gives infinitely many non-isomorphic algebras.

In the second case, there exists a basis x, y, z such that

$$[x, y] = y, \quad [x, z] = y + z, \quad [y, z] = 0. \quad (6.12.8)$$

- If $s = 3$, the algebra \mathfrak{g} must be simple. When \mathbb{K} is algebraically closed, there is only one such algebra, up to isomorphism, namely $\mathfrak{sl}_2(\mathbb{K})$. One can choose a basis of this algebra in such a way that

$$[x, y] = z, \quad [y, z] = x, \quad [z, x] = y. \quad (6.12.9)$$

However, over a field that isn't algebraically closed, this algebra may have several non-isomorphic forms. For example, over $\mathbb{K} = \mathbb{R}$ there are exactly two non-isomorphic 3-dimensional simple algebras, namely (\mathbb{R}^3, \times) and $\mathfrak{sl}_2(\mathbb{R})$.

Exercise 6.14. Show that all classical Lie algebras (that is, the algebras of matrices that preserve a non-degenerate bilinear form B in the sense that $BX + X^t B = 0$) over a field of characteristic 0 are simple with the exception of

$$\mathfrak{so}_4(\mathbb{K}) = \mathfrak{sl}_2(\mathbb{K}) \oplus \mathfrak{sl}_2(\mathbb{K}). \quad (6.12.10)$$

To prove this, first show that, for $\text{char}(\mathbb{K}) \neq 2$, the smallest ideal of $\mathfrak{sl}_n(\mathbb{K})$ containing any

standard unit e_{ij} is $\mathfrak{sl}_n(\mathbb{K})$ itself. Furthermore, letting $\text{char}(K) = p$ and assuming that $p \neq 2$ and p doesn't divide n , use this result to show that $\mathfrak{sl}_n(K)$ is simple for any $n \geq 2$.

This argument breaks down when p divides n : the trace of any scalar matrix (λI) vanishes in this case. Since scalar matrices are central, it follows that $\mathfrak{sl}_n(\mathbb{K})$ has a 1-dimensional ideal, and thus cannot be simple.

6.13 (*) Infinite-dimensional Lie groups

Many important groups, such as groups of diffeomorphisms, gauge groups, or loop groups, do not have the structure of a finite-dimensional manifold. However, they are still “locally Euclidean” in the sense that each point in them has a neighborhood that can be homeomorphically mapped onto a subset of an infinite-dimensional vector space. With a carefully defined smooth structure on that vector space, one can introduce the structure of an infinite-dimensional Lie group. However it is important to keep in mind that there is a huge variety of non-equivalent choices of topologies and smooth structures on infinite-dimensional spaces, so the particular definition will depend on the application. Here we will work with the most common class of infinite-dimensional Lie groups, which are locally modeled on Fréchet vector spaces.

Throughout this section we assume that \mathbb{K} is either \mathbb{R} or \mathbb{C} .

Definition 6.13.1 (Seminorm). A seminorm on a \mathbb{K} -vector space V is a function $p : V \rightarrow [0, \infty)$ which satisfies all properties of a norm except positive definiteness. That is, $p(x) = 0$ does not necessarily imply $x = 0$. We also say that p is *not point-separating*.

Definition 6.13.2 (Separating family of seminorms). Given a seminorm p on a \mathbb{K} -vector space V , let $N_p = \ker p = p^{-1}(0)$. Then the quotient $V_p = V/N_p$ is a normed space with $\|v + N_p\| = p(v)$. Denote by $\alpha_p : V \rightarrow V_p$ the corresponding quotient map.

A family \mathcal{P} of seminorms on V is called separating if

$$\forall p \in \mathcal{P} \quad p(v) = 0 \Rightarrow v = 0. \tag{6.13.1}$$

This is equivalent to the map

$$\alpha : V \rightarrow \prod_{p \in \mathcal{P}} V_p, \quad v \mapsto (\alpha_p(v))_{p \in \mathcal{P}}, \tag{6.13.2}$$

being injective.

Definition 6.13.3 (Locally convex topology). Recall that the initial topology defined on a set X by a family of mappings $f_j : X \rightarrow X_j$, $j \in J$, into topological spaces X_j is just the pre-image of the product topology under $f = \prod_{j \in J} f_j : X \rightarrow \prod_{j \in J} X_j$.

To each separating family \mathcal{P} of seminorms on V we associate the initial topology $\tau_{\mathcal{P}}$ on V

defined by the maps $\alpha_p : V \rightarrow V_p$ into the normed spaces V_p . We call it the locally convex topology on V defined by \mathcal{P} .

Since \mathcal{P} is separating, this topology makes V a Hausdorff space. It is also easy to show that V is a topological vector space in the sense that addition and multiplication by scalars on V are continuous maps. A topological vector space $(V, \tau_{\mathcal{P}})$ constructed this way is called a *locally convex space*. As follows from the construction, each locally convex space is in particular a topological vector space which can be embedded into a product $\prod_{p \in \mathcal{P}} V_p$ of normed spaces.

Definition 6.13.4. Some authors drop the requirement that \mathcal{P} be separating, in which case the resulting topology is not Hausdorff.

Definition 6.13.5 (Fréchet space). A locally convex space V is called a Fréchet space if its topology can be defined by a countable family $\mathcal{P} = \{p_n\}_{n=0}^{\infty}$ of seminorms and if V is complete with respect to the compatible metric

$$d(x, y) = \sum_{n=0}^{\infty} 2^{-n} \frac{p_n(x - y)}{1 + p_n(x - y)}. \quad (6.13.3)$$

This is equivalent to V being complete with respect to $\{p_n\}_n$ (i.e. if a sequence $(x_i)_i \in V$ is Cauchy with respect to each seminorm p_n , then there is a $x \in V$ such that $\lim_{i \rightarrow \infty} p_n(x_i) = 0$ for each n ; equivalently, $\alpha_{p_n}(x_i)$ is a Cauchy sequence in V_p for each n).

Remark 6.13.6. The introduction of a metric is not necessary for defining Fréchet spaces. In fact, a Fréchet space is simply a locally convex metrizable topological vector space that is complete as a topological vector space. Note that a metrizable space is complete¹⁸ iff it is sequentially complete. The latter means that every Cauchy sequence has a limit, and a Cauchy sequence in a topological vector space is a sequence $(x_n)_n \in V$ such that for every open neighborhood $U \subset V$ of the origin, $x_n - x_m \in U$ for sufficiently large n and m . This definition doesn't rely on the existence of any metric. However, since it is easy to check that $d(x, y)$ based on a separating family of seminorms $\{p_n\}$ defines a translation invariant metric on any locally convex space, it follows that these two definitions are equivalent.

Moreover, Fréchet spaces can also be defined as topological vector spaces whose topology can be induced by some translation-invariant metric.

Example 6.13.7. (a) Let X be a topological space. For each compact $K \subset X$ we obtain a seminorm p_K on $C(X, \mathbb{R})$ by

$$p_K(f) = \sup_{x \in K} |f(x)|. \quad (6.13.4)$$

The family \mathcal{P} of these seminorms defines on $C(X, \mathbb{R})$ the locally convex topology of uniform convergence on compact subsets of X .

In general, this is not a metrizable and thus not a Fréchet space. However if X is compact, then we may take $K = X$ and obtain a norm on $C(X, \mathbb{R})$, making all other seminorms redundant. In this case $C(X, \mathbb{R})$ is a Banach space (a complete normed vector space).

- (b) The preceding example can be generalized to the space $C(X, V)$, were V is a locally convex space. We define for each compact $K \subset X$ and each continuous seminorm q on V a seminorm

$$p_{K,q}(f) = \sup_{x \in K} q(f(x)). \quad (6.13.5)$$

The family of these seminorms defines a locally convex topology on $C(X, V)$. In fact, it is not difficult to check that this topology coincides with the compact-open topology.

- (c) If X is locally compact and countable at infinity,¹⁹ then there exists a sequence K_n of compact subsets of X with $\bigcup_n K_n$ and $K_n \subset \text{int } K_{n+1}$. We call such a sequence an *exhaustion* of X . Then each compact subset $K \subset X$ lies in some K_n , so that each seminorm p_K is dominated by some p_{K_n} . This implies that $C(X, \mathbb{R})$ is metrizable (because the metric $d(x, y)$ can be defined in this case), and since it is also complete, it is a Fréchet space. It even is a Fréchet algebra in the sense that the multiplication is continuous.
- (d) For any set X the space \mathbb{R}^X of all real-valued functions $X \rightarrow \mathbb{R} = \prod_{y \in \mathbb{R}} X$ is a locally convex space with respect to the product topology. This topology is defined by the seminorms $p_x(f) = |f(x)|$, $x \in X$. This space is complete, and it is metrizable iff X is countable.

Example 6.13.8. (a) Let $U \subset \mathbb{R}^n$ be an open subset and consider the algebra $C^\infty(U, \mathbb{R})$. For each multiindex $m = (m_1, \dots, m_n) \in \mathbb{N}_0^n$ with $|m| = m_1 + \dots + m_n$, we consider the differential operator

$$D^m = D_1^{m_1} \cdots D_n^{m_n} = \frac{\partial^{|m|}}{\partial_1^{m_1} \cdots \partial_n^{m_n}}. \quad (6.13.6)$$

For each m and each compact subset $K \subset U$ we define a seminorm on $C^\infty(U, \mathbb{R})$ by

$$p_{K,m}(f) = \sup_{x \in K} |D^m f(x)|. \quad (6.13.7)$$

The family of these seminorms defines a locally convex topology on $C^\infty(U, \mathbb{R})$.

To obtain an exhaustion of U , we choose a norm $\|\cdot\|$ on \mathbb{R}^n and consider the compact subsets

$$K_n = \{x \in U \mid \|x\| \leq n, \text{dist}(x, \mathbb{R}^n \setminus U) \geq 1/n\}, \quad (6.13.8)$$

where $\text{dist}(x, \mathbb{R}^n \setminus U) = \inf_{y \in \mathbb{R}^n \setminus U} \|x - y\|$ is a continuous function. It is easy to see that $(K_n)_n$ is an exhaustion of U , so that the topology on $C^\infty(U, \mathbb{R})$ can be defined by

a countable set of seminorms. Moreover, $C^\infty(U, \mathbb{R})$ is complete within respect to the corresponding metric and the multiplication on this algebra is continuous, so it is a Fréchet algebra.

- (b) Let M be a smooth n -dimensional manifold and consider the algebra $C^\infty(M, \mathbb{R})$. If (U, φ) is a chart of M , then $\varphi(U)$ is an open subset of \mathbb{R}^n , so that, in view of (a), we have a Fréchet algebra structure on $C^\infty(\varphi(U), \mathbb{R})$. We now consider the map

$$\Phi : C^\infty(M, \mathbb{R}) \hookrightarrow \prod_{\alpha} C^\infty(\varphi_\alpha(U_\alpha), \mathbb{R}), \quad f \mapsto (f|_U \circ \varphi^{-1})_\alpha \quad (6.13.9)$$

where α is an index enumerating all charts $(U_\alpha, \varphi_\alpha)$ in an atlas of M , and endow the right hand side with the product topology, turning it into a locally convex algebra (this is an easy exercise). Therefore the inverse image of this topology turns $C^\infty(M, \mathbb{R})$ into a locally convex algebra.

This description is convenient but not very explicit. To see how it can be defined by seminorms, note that for each compact $K \subset M$ for which there exists a chart $\varphi : U \rightarrow \mathbb{R}^n$ with $K \subset U$ and for each multiindex $m \in \mathbb{N}_0^n$ we have a seminorm

$$p_{K,m}(f) = \sum_{x \in \varphi(K)} |D^m(f \circ \varphi^{-1})|. \quad (6.13.10)$$

It is easy to see that these seminorms define the same topology on $C^\infty(M, \mathbb{R})$ and that we obtain the structure of a Fréchet algebra on $C^\infty(\mathbb{R}, \mathbb{R})$. This is called the *topology of local uniform convergence of all partial derivatives*.

- (c) If M is a finite-dimensional paracompact complex manifold, then we consider the algebra $\text{Hol}(M, \mathbb{C})$ of holomorphic functions on M as a subalgebra of $C(M, \mathbb{C})$, endowed with the topology of uniform convergence on compact subsets of M . This topology turns $\text{Hol}(M, \mathbb{C})$ into a Fréchet algebra. Moreover, one can show that the injective map $\text{Hol}(M, \mathbb{C}) \hookrightarrow C^\infty(M, \mathbb{C})$ is also a topological embedding.

Definition 6.13.9 (Final locally convex topology). Given a family of linear maps $f_\alpha : V_\alpha \rightarrow V$ from locally convex spaces V_α to a vector space V , the final locally convex topology on V for this family is the locally convex topology defined by all those seminorms p on V for which all compositions $p \circ f_\alpha$ are continuous seminorms on V_α .

This topology has the universal property that a linear map $\varphi : V \rightarrow W$ into a locally convex space W is continuous iff all the maps $\varphi \circ f_\alpha$ are continuous.

Example 6.13.10. (a) Let X be a locally compact space and $C_c(X, \mathbb{R})$ the space of compactly supported continuous functions. For each compact subset $K \subset X$ we then have

the natural inclusion

$$\alpha_K : C_K(X, \mathbb{R}) \hookrightarrow C_c(X, \mathbb{R}), \quad (6.13.11)$$

where $C_K(X, \mathbb{R})$ is the subspace consisting of functions supported on K . Each space $C_K(X, \mathbb{R})$ is a Banach space with respect to the sup-norm $\|f\|_\infty$. We then endow $C_c(X, \mathbb{R})$ with the final locally convex topology defined by the maps α_K .

- (b) Let M be a smooth manifold and consider the space $C_c^\infty(M, \mathbb{R})$ of compactly supported smooth functions. For each compact $K \subset M$ we have a natural inclusion

$$\alpha_K : C_K^\infty(M, \mathbb{R}) \hookrightarrow C_c^\infty(M, \mathbb{R}). \quad (6.13.12)$$

we endow each space $C_K^\infty(M, \mathbb{R})$ with the subspace topology inherited from $C^\infty(M, \mathbb{R})$, which turns it into a Fréchet space. We then endow $C_c^\infty(M, \mathbb{R})$ with the final locally convex topology defined by the maps α_K .

Definition 6.13.11 (Calculus on locally convex spaces). Let X, Y be topological vector spaces, $U \subset X$ open and $f : U \rightarrow Y$ a map. Then the *directional derivative* of f with respect to $h \in X$ is

$$df(x)(h) = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + th) - f(x)) \quad (6.13.13)$$

whenever the limit exists. The function f is called *differentiable* if $df(x)(h)$ exists for all $h \in X$, It is called *continuously differentiable* if it is differentiable at all points of U and

$$df : U \times X \rightarrow Y, \quad (x, h) \mapsto df(x)(h) \quad (6.13.14)$$

is a continuous map. It is called a C^1 -map if it is continuous and continuously differentiable, a C^n -map if df is a C^{n-1} -map, and a C^∞ - (or smooth) map if it is C^n for all $n \geq 1$.

if X and Y are complex vector spaces, then the map f is called *holomorphic* if it is C^1 and for all $x \in U$ the map $df(x) : X \rightarrow Y$ is complex linear (we will show that df is always real linear).

Remark 6.13.12. The existence of linear maps (which are always differentiable in the above sense) that are not continuous in the infinite-dimensional case shows that continuity of f does not follow from its differentiability.

Lemma 6.13.13. Let X, Y be locally convex spaces, $U \subset X$ open and $f : U \rightarrow Y$ a continuously differentiable map. Then:

- (i) For any $x \in U$ the map $df(x) : X \rightarrow Y$ is real linear and continuous.

(ii) (*Fundamental Theorem of Calculus*) If $x + [0, 1]h \subset U$, then

$$f(x + h) = f(x) + \int_0^1 df(x + th)(h)dt. \quad (6.13.15)$$

In particular, f is locally constant iff $df = 0$.

(iii) f is continuous.

(iv) If f is C^n , $n \geq 2$, then the functions $(h_1, \dots, h_n) \mapsto d^n f(x)(h_1, \dots, h_n)$, $x \in U$, are symmetric n -linear maps.

(v) If $x + [0, 1]h \subset U$, then we have the Taylor formula

$$\begin{aligned} f(x + h) &= f(x) + df(x)(h) + \dots + \frac{1}{(n-1)!} d^{n-1} f(x)(h, \dots, h) + \\ &\quad + \frac{1}{(n-1)!} \int_0^1 (1-t)^{n-1} d^n f(x + th)(h, \dots, h) dt. \end{aligned} \quad (6.13.16)$$

Proof. (i) For each linear functional $\lambda \in Y^*$ and $h_1, h_2 \in X$ the map

$$F(t_1, t_2) = \lambda(f(x + t_1 h_1 + t_2 h_2)) \quad (6.13.17)$$

is defined on an open neighborhood of the origin in \mathbb{R}^2 and has continuous partial derivatives

$$\partial_{t_i} F = df(x + t_1 h_1 + t_2 h_2)(h_i), \quad i = 1, 2. \quad (6.13.18)$$

From finite-dimensional calculus we know that F is a C^1 -map and $dF(0, 0) : \mathbb{R}^2 \rightarrow \mathbb{R}$ is linear. This implies that $\lambda \circ df(x)$ is linear on $\text{Span}(h_1, h_2)$. Since Y^* separates the points of Y (i.e. the pairing with the dual space is non-degenerate by definition) and h_1, h_2 are arbitrary, the map $df(x)$ is real linear. Its continuity follows from the continuity of df .

(ii) We consider for $\lambda \in Y^*$ the C^1 -map

$$F : I \rightarrow \mathbb{R}, \quad F(t) = \lambda(f(x + th)) \quad (6.13.19)$$

and obtain from the Fundamental Theorem in single-variable calculus

$$\lambda(f(x+h) - f(x)) = F(1) - F(0) = \int_0^1 F'(t)dt = \int_0^1 \lambda(df(x+th)(h))dt. \quad (6.13.20)$$

Since Y^* separates the points of Y , this implies that the weak integral $\int_0^1 df(x+th)(h)dt$, which a priori exists only in the completion of Y , actually defines an element of Y which coincides with $f(x+h) - f(x)$.

- (iii) Let p be a continuous seminorm on Y and $\epsilon > 0$. Then there exists a balanced neighborhood (i.e. one that is invariant under multiplications by scalars a such that $|a| \leq 1$) of the origin $U_1 \subset X$ with $x + U_1 \subset U$ and $p(df(x + th)(h)) < \epsilon$ for $t \in [0, 1]$ and $h \in U_1$. Hence

$$p(f(x + h) - f(x)) \leq \int_0^1 p(df(x + th)(h)) dt \leq \epsilon \quad (6.13.21)$$

(this is an easy exercise), and this f is continuous.

- (iv) Arguing as in (i), we may WLOG assume that $Y = \mathbb{R}$. That the maps $d^n f(x)$ are symmetric and linear follows by considering maps of the form

$$(t_1, \dots, t_n) \mapsto f(x + t_1 h_1 + \dots + t_n h_n) \quad (6.13.22)$$

on an open neighborhood of the origin in \mathbb{R}^n and then applying the corresponding finite-dimensional result.

- (v) We consider the C^n -map

$$F : I \rightarrow \mathbb{R}, \quad F(t) = f(x + th), \quad F^{(n)}(t) = d^n f(x + th)(h, \dots, h) \quad (6.13.23)$$

and apply the Taylor formula for C^n -functions $I \rightarrow \mathbb{R}$.

■

From these properties, one can similarly derive the rest of the classic theorems of calculus, such as the chain rule

$$d(g \circ f)(x) = dg(f(x)) \circ df(x), \quad (6.13.24)$$

and the equivalence of the C^1 property on the product space $X_1 \times X_2$ to the continuity of partial derivatives $d_1 f(x_1, x_2)$ and $d_2 f(x_1, x_2)$. We refer the reader to [34, Sec. II] for details.

Remark 6.13.14. The following properties immediately follows from these theorems.

- (a) Continuous linear maps between locally convex spaces are smooth.
- (b) The addition map $+ : X \times X \rightarrow X$ on topological vector spaces is smooth.
- (c) Continuous k -linear maps $m : X_1 \times \dots \times X_k \rightarrow Y$ are smooth.
- (d) Assuming $f \in C^n$, f is C^{n+1} iff $d^n f$ is C^1 .
- (e) If $f : U \rightarrow Y$ is holomorphic, then $df : U \times X \rightarrow Y$ is also holomorphic (since from finite-dimensional theory each map $x \mapsto df(x)(h)$ is holomorphic).

Remark 6.13.15. The assumption of local convexity in the definition of C^1 -maps is crucial for the validity of the Fundamental Theorem. Without local convexity, it is possible to construct non-constant “ C^1 ” paths $\gamma : I \rightarrow X$ such that $d\gamma = 0$.

In the context of Banach spaces, one further has an Inverse Mapping Theorem and an Implicit

Mapping Theorem. Such results cannot be expected in general for Fréchet spaces (e.g. the exponential function on the group $\text{Diff}(\mathbb{S}^1)$ of diffeomorphisms of the circle). Nevertheless, there is an Implicit Mapping Theorem for maps of the type $E \rightarrow F$, where F is Banach and E is locally convex.

Remark 6.13.16 (Pathologies of linear ODEs in Fréchet spaces). (a) Consider the Fréchet space $V = C^\infty([0, 1], \mathbb{R})$ and the continuous linear operator $Lf = f'$ on V . We are asking for solutions of the initial value problem

$$\gamma'(t) = L\gamma(t), \quad \gamma(0) = \gamma_0 \in V. \quad (6.13.25)$$

As a consequence of Borel's Theorem that each power series is the Taylor series of some smooth function, each γ_0 has a smooth extension to a function on \mathbb{R} . Let h be such a function and consider

$$\gamma : \mathbb{R} \rightarrow V, \quad \gamma(t)(x) = h(t + x). \quad (6.13.26)$$

Then $\gamma(0) = h|_{[0,1]} = \gamma_0$ and $\gamma'(t)(x) = h'(t + x) = (L\gamma(t))(x)$. It is clear that these solutions depend on the choice of the extension h . Therefore solutions to such initial problem values are not unique in Fréchet spaces, unlike in the finite-dimensional case.

- (b) Now we consider $V = C^\infty(\mathbb{S}^1, \mathbb{C})$ identified with the space of 2π -periodic smooth functions on the real line. We consider the linear operator $Lf = -f''$ and the same equation (6.13.25), which in this case is the heat equation with reversed time. It is easy to analyze this equation in terms of the Fourier expansion of γ . Let

$$\gamma(t)(x) = \sum_{n \in \mathbb{Z}} a_n(t) e^{inx} \quad (6.13.27)$$

be the Fourier series for $\gamma(t)$. Then equation (6.13.25) implies $a'_n(t) = n^2 a_n(t)$ for each $n \in \mathbb{Z}$, so, $a_n(t) = a_n(0) e^{\epsilon n^2 t}$ for any solution γ . If the Fourier coefficients $a_n(0)$ of γ do not satisfy

$$\sum_n |a_n(0)| e^{\epsilon n^2} < \infty \quad (6.13.28)$$

for some $\epsilon > 0$ (which need not be the case for a smooth function γ_0), then the initial value problem (6.13.25) has no solutions on $[0, \epsilon]$.

As a consequence, the operator $\exp(L)$ is never defined for $t \neq 0$. Nevertheless, we may use the Fourier series to see that $\beta(t) = (1 + it^2)I + tL$ (where I is the identity operator) defines a curve $\beta : \mathbb{R} \rightarrow \text{GL}(V)$ which is smooth in the sense that

$$\mathbb{R} \times V \rightarrow V \times V, \quad (t, v) \mapsto (\beta(t)(v), \beta(t)^{-1}(v)) \quad (6.13.29)$$

is smooth. We further have $\beta'(0) = L$, so that L arises as the tangent vector of a smooth curve in $\text{GL}(V)$, but not for a one-parameter group.

Definition 6.13.17 (Mackey complete space). A locally convex space E is called Mackey complete if for each sooth curve $\xi : [0, 1] \rightarrow E$ there exists a smooth curve $\eta : [0, 1] \rightarrow E$ with $\eta' = \xi$.

Remark 6.13.18. If E is a sequentially complete locally convex space (e.g. a Fréchet space), then it is Mackey complete because sequential completeness implies the existence of Riemann integrals of continuous E -valued functions on compact intervals, hence that we can put $\eta(t) = \int_0^t \xi(s)ds$ and then $\eta' = \xi$.

Definition 6.13.19 (Infinite-dimensional smooth manifold). Let M be a Hausdorff topological space and E a locally convex space. An E -chart of an open subset $U \subset M$ is a homeomorphism $\varphi : U \rightarrow \varphi(U) \subset E$ into an open subset $\varphi(U)$ of E . Two charts $(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) are called smoothly compatible if the map

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_{\alpha\beta}) \rightarrow \varphi_\beta(U_{\alpha\beta}) \quad (6.13.30)$$

is smooth. From the chain rule it follows that compatibility of charts is an equivalence relation on the set of all E -charts. An E -atlas is a family of pairwise compatible E -charts that cover all of M . A smooth E -structure \mathcal{A} on M is a maximal E -atlas. A smooth E -manifold is a pair (M, \mathcal{A}) where \mathcal{A} is a smooth E -structure on M .

We call a manifold modeled on a locally convex, resp., Fréchet space, resp., Banach space a locally convex, resp., Fréchet, resp., Banach manifold.

All of differential geometry that we will cover for finite-dimensional smooth manifolds translates without changes to the case of manifolds modeled on locally convex spaces. This includes tensor bundles, differential forms, Lie derivatives, exterior derivatives, and formulas relating them. However, results involving integration, such as the definition of integrals of differential forms, Stokes' theorem, and Poincaré lemma, require Mackey completeness.

Definition 6.13.20 (Locally convex Lie group). A locally convex Lie group is a locally convex manifold endowed with a group structure such that the multiplication and the inversion maps are smooth.

- Example 6.13.21.**
- (a) (Vector groups) Each locally convex space V is an abelian Lie group with respect to addition.
 - (b) (Diffeomorphism groups) For a compact finite-dimensional manifold M , the group $\text{Diff}(M)$ of its diffeomorphisms can be turned into a locally convex Lie group modeled on the space $\mathfrak{X}(M)$ of smooth vector fields on M . Since we have not yet defined this space, the details of example will have to wait, but it will be useful to keep in mind for some of the discussions.

7 Fiber Bundles

7.1 Bundles

Recall that by the Rank Theorem 3.7.1, all submersions when restricted to sufficiently small neighborhoods in the domain look like projections in a direct product. Fiber bundles are a natural special case of this, where the submersion looks like a projection “locally in the target”, i.e. is a locally trivial map.

Definition 7.1.1 (Topological fiber bundle). A topological fiber bundle is continuous locally trivial map $\pi : E \rightarrow B$. The local trivializations $\varphi_\alpha : \pi^{-1}(U_\alpha) \rightarrow U_\alpha \times F$ are called *bundle chart maps*. The set of those $b \in B$ for which the *fiber* $\pi^{-1}(b)$ is homeomorphic to a fixed space F is open and closed in B . Therefore it usually suffices to fix the homeomorphism type of the fibers. If all fibers are homeomorphic to F , we call F the *typical, or model, fiber*.

From this definition it is clear that topological covering maps are exactly the topological fiber bundles with discrete fibers. Also by Theorem 5.9.3, every fiber bundle is a fibration, which we state as a Corollary.

Corollary 7.1.2 ([17, Cor. 3.2.5]). *Topological fiber bundles are Serre fibrations.*

This implies that all of our topological results for Serre fibrations will also hold for fiber bundles, most importantly the long exact sequence of homotopy groups. If the base space happens to be paracompact, one has the following stronger result.

Corollary 7.1.3 (Huebsch and Hurewicz [17, Prop. 3.2.6]). *Topological fiber bundles over paracompact Hausdorff base spaces are Hurewicz fibrations.*

Proof. See [26, §13.4]. ■

This will property will obviously survive in the smooth category since all manifolds are paracompact.

Definition 7.1.4 (Smooth fiber bundle/locally trivial fibration). A map $\pi \in C^\infty(E, M)$ is a locally trivial fibration, or a **fiber bundle (FB)**, if it is a smooth submersion and for any $p \in M$ there is an open neighborhood U_p and a diffeomorphism $\chi_p : \pi^{-1}(U_p) \rightarrow U_p \times \pi^{-1}(p)$ (recall that $\pi^{-1}(p)$ is a smooth manifold for submersions) such that the following triangle

commutes:

$$\begin{array}{ccc}
 \pi^{-1}(U_p) & \xrightarrow{\chi_p} & U_p \times \pi^{-1}(p) \\
 \pi|_{\pi^{-1}(U_p)} \searrow & \curvearrowright & \swarrow \text{pr}_1 \\
 & U_p &
 \end{array} \tag{7.1.1}$$

It is called a *trivial bundle* if U can be taken to be all of M , i.e. the total space decomposes into a product $E \cong M \times \pi^{-1}(p)$. The maps χ_p are called *local trivializations*.

Finally, a **FB** can be restricted to an open subset $U \subset M$ by defining $E|_U$ as $\pi^{-1}(U) \xrightarrow{\pi|_{\pi^{-1}(U)}} M$.

Exercise 7.1. Suppose $E \xrightarrow{\pi} M$ is a **FB** with fiber F . Prove the following:

- (a) π is an open quotient map.
- (b) E is compact iff both M and F are compact.

One is tempted to say that the **HLP** property of automatically applies to smooth fiber bundles and stop at that. However, in the smooth category we want the liftings of smooth homotopies to also be smooth, and unfortunately the topological versions of these theorems will not suffice for that. This is why we will have to enhance many of the familiar results about fibrations for the smooth case. We start by showing that all fibers in a fiber bundle are not just homotopy equivalent, but diffeomorphic (similarly, all fibers in a topological fiber bundle are homeomorphic).

Proposition 7.1.5. *For a **FB** $\pi : E \rightarrow M$ with M connected, all fibers $\pi^{-1}(m)$ are diffeomorphic to each other. We call a manifold F that they are all diffeomorphic to “the typical, or model, fiber of the fibration”.*

Proof. A connected manifold is path-connected. By connecting any two points with a path, covering this path by a finite sequence of open sets over which the fibration is trivial (which is possible by compactness of $[0, 1]$), and using the chain of local trivializations, we establish a diffeomorphism between the fibers at the endpoints. ■

Example 7.1.6. 1. The cylinder $E = \mathbb{S}^1 \times \mathbb{R}$ is a trivial *line bundle* (or rather the total space of one) over a circle with π the projection onto the first factor. It is also a *circle bundle* over the line under the projection onto the second factor.
 2. The Möbius band is a non-trivial line bundle over the circle under the projection onto the “equator” of the band. It is not trivial because it’s not homeomorphic to the trivial line bundle, which is the cylinder.

3. $\mathbb{R}^n \setminus \{0\}$ is a trivial line bundle above the sphere \mathbb{S}^{n-1} under $\pi(x) = x/\|x\|$.

Example 7.1.7. Covering spaces of M are exactly the **FBs** over M whose fibers are at most countable sets with discrete topology.

The following fundamental theorem gives us a useful sufficient condition for a map to be a locally trivial fibration. We delay its proof until Section 9.2 because it requires some theory of flows of vector fields. Nevertheless we present it now because it will be useful for constructing specific examples of fiber bundles. Recall that a map is proper iff it is closed and its fibers (preimages of points) are compact.

Theorem 7.1.8 (Ehresmann fibration theorem). *If $\pi : E \rightarrow M$ is a proper surjective submersion, then π is a locally trivial fibration.*

Corollary 7.1.9. *A surjective submersion acting between compact manifolds is a locally trivial fibration.*

Example 7.1.10 (Qubit/Bloch sphere/Hopf fibration). A qubit is a two-level system described by a pair of complex numbers $(z_1, z_2) \in \mathbb{C}^2$ such that $|z_1|^2 + |z_2|^2 = 1$. The set of such points is the unit sphere \mathbb{S}^3 in $\mathbb{C}^2 \cong \mathbb{R}^4$. However, points differing only by a phase correspond to the same physical state: $(e^{i\theta} z_1, e^{i\theta} z_2) \sim (z_1, z_2)$. It turns out that these equivalence classes are in one-to-one correspondence with points of a two-sphere \mathbb{S}^2 , because the quotient map can be given by

$$\pi(z_1, z_2) = (\underbrace{2z_1\bar{z}_2}_{\in \mathbb{D} \subset \mathbb{C}}, \underbrace{|z_1|^2 - |z_2|^2}_{\in [-1, 1]}) \in \mathbb{S}^2 \subset \mathbb{R}^3 \cong \mathbb{C} \times \mathbb{R}. \quad (7.1.2)$$

This is obviously a smooth and surjective map, and since the sphere \mathbb{S}^3 is compact, it is proper. It is also easy to check that its differential is everywhere surjective, therefore by Ehresmann's theorem π is a **FB** with typical fiber \mathbb{S}^1 (space of phases). This fibration, often written as $\mathbb{S}^3 \xrightarrow{\$^1} \mathbb{S}^2$, is called the Hopf fibration, or the Hopf bundle. Since \mathbb{S}^3 with one point excluded is homeomorphic to \mathbb{R}^3 , the Hopf fibration is often visualized as a fibration of the 3D space by circles plus one straight line (circle of infinite radius). Any two circles in this fibration happen to be linked with each other, see Fig. 7.1.

\mathbb{S}^2 , which represents the set of physical states of a qubit, is the Bloch sphere. One can show that this fibration is not trivial (the easiest way, perhaps, is to compare the fundamental groups $\pi_1(\mathbb{S}^2 \times \mathbb{S}^1) = \pi_1(\mathbb{S}^2) \times \pi_1(\mathbb{S}^1) = \mathbb{Z}$ and $\pi_1(\mathbb{S}^3) = \{e\}$).

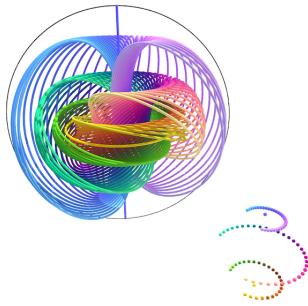


Fig. 7.1: Hopf fibration in stereographic projection, i.e. as a fibration of \mathbb{R}^3 . Each circle is a fiber corresponding to a point on the 2-sphere at the bottom. From [15].

Exercise 7.2. Show that in the Hopf bundle, the pre-image of any circle in \mathbb{S}^2 is a two-torus \mathbb{T}^2 embedded into \mathbb{S}^3 .

Definition 7.1.11 (Bundle morphism). Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be two smooth **FBS**s. A bundle morphism (a.k.a. *bundle map*) between them is a pair of smooth maps $\tilde{h} : E \rightarrow E'$ and $h : M \rightarrow M'$ such that \tilde{h} “covers” h in the sense that the following square commutes:

$$\begin{array}{ccc} E & \xrightarrow{\tilde{h}} & E' \\ \pi \downarrow & \curvearrowright & \downarrow \pi' \\ M & \xrightarrow{h} & M' \end{array} \quad (7.1.3)$$

This gives rise to the category Bun^∞ of all smooth **FBS**s. A full subcategory is the category Bun_M^∞ of all smooth **FBS**s over a given base M with $h = \text{id}_M$ (note that in general there are other bundle morphisms between bundles over M for which h is not even a diffeomorphism). Bundle morphisms covering the identity are also called *vertical*.

7.2 Structure groups

So far we have only given a purely topological (or geometric in the smooth case) description of fiber bundles. The fascinating thing about fiber bundles, however, is that the requirement of local triviality gives rise to an alternative *algebraic* description of fiber bundles, which will eventually lead us into the realm of cohomology. A similar, but more cumbersome, description appears for general topological fibrations as well. We now characterize fiber bundles in terms of their systems of local trivializations and identify their transition maps

with certain Lie group actions. Note that everything here easily translates to the category of topological fiber bundles as well.

Definition 7.2.1 (Bundle atlas, G -structure, cocycle). Let $E \xrightarrow{\pi} M$ be a smooth **FB** with typical fiber F , and let $\{U_\alpha\}_\alpha$ be an open covering of M such that the restrictions $E|_{U_\alpha}$ are trivial. By definition of local triviality, we have some diffeomorphisms $\chi_\alpha : E|_{U_\alpha} \rightarrow U_\alpha \times F$, called *local trivializations*. Then:

- The collection $\{(U_\alpha, \chi_\alpha)\}$ is called a **FB atlas** for $E \xrightarrow{\pi} M$. Two bundle atlases are *equivalent* if their union is still a bundle atlas. A *bundle structure* on E is an equivalence class of bundle atlases, or, alternatively, a *maximal atlas* given by the union of all atlases belonging to an equivalence class.
- For any point $(m, f) \in U_{\alpha\beta} \times F$, the transition maps in this atlas have the form

$$\chi_\beta \circ \chi_\alpha^{-1}(m, f) = (m, t_{\beta\alpha}(m, f)), \quad (7.2.1)$$

, where for every fixed $m \in M$, the map $t_{\beta\alpha}(m, _) : F \rightarrow F$ is a diffeomorphism. One can think of $t_{\beta\alpha}$ as a function on $U_{\alpha\beta}$ that takes values in the group $\text{Diff}(F)$ of all diffeomorphisms of F , so $t_{\beta\alpha}$ becomes a map $t_{\beta\alpha} : U_{\alpha\beta} \rightarrow \text{Diff}(F)$.²⁰ These $\text{Diff}(F)$ -valued functions are called *bundle transition functions*. They are also sometimes called *clutching functions*, especially when M is a sphere. We will usually assume that $\text{Diff}(F)$ acts on the typical fiber from the *left* (one can pass to a right action by acting via $t_{\alpha\beta}^{-1}$ instead).

- If the bundle structure of E contains a bundle atlas whose transition functions take values in a Lie subgroup $G < \text{Diff}(F)$,²¹ then such an atlas is called a *G-atlas* for E . Two G -atlases are called *G-equivalent* if their union is still a G -atlas. A G -equivalence class of G -atlases, or, alternatively, the maximal G -atlas equal to their union, is called a *G-structure*, denoted \mathcal{G} . If a G -structure is fixed on E , we say that G is the *structure group* of the bundle. The tuple $(E \xrightarrow{\pi} M, G \curvearrowright F, \mathcal{G})$ is called a *G-bundle*.
- If $H < G$ is a Lie subgroup and a given G -structure on E contains an H -atlas, then a choice of an H -structure (which is a subset of the given G -structure) is called a *reduction of the structure group*, or simply a *bundle reduction* from G to H .
- From the definition (7.2.1) it is obvious that the group-valued bundle transition functions satisfy two important conditions:

$$\begin{cases} t_{\alpha\alpha} \equiv e, & (\text{identity condition}), \\ t_{\alpha\beta} \cdot t_{\beta\gamma} \cdot t_{\gamma\alpha}|_{U_{\alpha\beta\gamma}} \equiv e; & (\text{cocycle condition}), \end{cases} \quad (7.2.2)$$

where the product in $\text{Diff}(F)$ is just the composition of maps. Any collection of such G -valued functions $\{t_{\alpha\beta}\}$ associated to an open covering of M is called a *cocycle* on M with structure group G .²² Note that these two conditions imply $t_{\alpha\beta}t_{\beta\alpha} = e$.

- Remark 7.2.2.**
1. One can try writing down higher-order relations on the transition functions, such as $t_{\alpha\beta}t_{\beta\gamma}t_{\gamma\delta}t_{\delta\alpha} = e$, but all of them follow from the identity and the cocycle conditions. Meanwhile, replacing the cocycle condition with only $t_{\alpha\beta}t_{\beta\gamma} = e$ would've been insufficient because this doesn't imply the higher-order relations.
 2. $\text{Diff}(F)$ acts on F faithfully by definition, therefore the induced action of any subgroup $G < \text{Diff}(F)$ is faithful as well. This is why in the following definitions we will assume a priori that the action of the structure group on the model fiber is faithful. We will further elucidate the importance of faithfulness in Remark 7.4.10.
 3. A fiber bundle $E \rightarrow M$ is trivial iff it admits an $\{e\}$ -structure. Indeed, if there is a bundle atlas in which all transition functions evaluate to the identity, then all local trivializations agree on the overlaps and hence stitch into a global trivialization.

Since the structure group of a bundle acts on its typical fiber F , that is, it acts in the local trivializations, it is natural to ask whether this action “lifts” to an action on the actual fibers of the bundle itself. Crucially, the answer is no, however the following restricted concept of a G -structure does get naturally translated to the fibers.

Definition 7.2.3 (G -fiber). Let F be a smooth manifold diffeomorphic to a given manifold F_0 and let G be a Lie group acting faithfully from the left on F_0 , so G can be identified with a Lie subgroup of $\text{Diff}(F_0)$. Two diffeomorphisms $\chi_1, \chi_2 : F \rightarrow F_0$ are called G -equivalent if $\chi_2 \circ \chi_1^{-1} \in G$. A G -equivalence class \mathcal{G}_0 of such diffeomorphisms, i.e. an orbit of the action of G on $\text{Diff}(F, F_0)$ by compositions $\chi \mapsto g \circ \chi$, is called a fiber G -structure on F .²³ The tuple $(F, G \curvearrowright F_0, \mathcal{G}_0)$ is called a G -fiber modeled on F_0 . An element of \mathcal{G}_0 is called a G -frame, or a frame compatible with the G -structure.

Proposition 7.2.4. *If $\pi : E \rightarrow M$ is a fiber bundle with structure group G acting on the typical fiber F , then each fiber of E naturally carries a fiber G -structure and an action of the center $Z(G)$.*

Proof. Consider a fiber $E_m = \pi^{-1}(m)$, $m \in M$. The fact that it carries a fiber G -structure is obvious because the value of any G -valued transition function at a point m is an element of G .

We now show that if G is not abelian, then there is no natural action of G on the fibers, but there is an action of $Z(G)$. By definition of the typical fiber, there is a diffeomorphism $f_\alpha = \chi_\alpha|_{E_m} : E_m \rightarrow F$, which is just the restriction of a local trivialization χ_α . By definition of a G -structure, there is a smooth (say left) action of G on the typical fiber F . Let

$f_\beta : E_m \rightarrow F$ be the diffeomorphism coming analogously from another compatible local chart. Then $f_\alpha(p) = t^{-1} \cdot f_\beta(p)$ for all $p \in E_m$ for some $t \in G$, and we can compare what the action of an element $g \in G$ translates to in the fiber E_m under these two diffeomorphisms:

$$f_\alpha^{-1}(g \cdot f_\alpha(p)) = f_\beta^{-1}(t g t^{-1} f_\beta(p)). \quad (7.2.3)$$

Since t can range over all of G by varying the charts, these formulas consistently define an action $g \cdot p$ only for $g \in Z(G)$. ■

- Remark 7.2.5.**
1. The intuition behind the definition of G -structures is the following. A fiber G -structure on F is a choice of a G -invariant subset of all diffeomorphisms from F to a “standard” or “model” space F_0 , namely an orbit of the action of G on all such diffeomorphisms by compositions. This can be interpreted as F_0 having some kind of “structure” that is invariant under actions of G , and a fiber G -structure allows us to transplant this structure to F . A G -structure on a fiber bundle, then, is a collection of structures on each fiber that, in a sense, vary smoothly with the base point $m \in M$. These structures are defined consistently because all local trivializations belonging to the same G -bundle atlas differ from each other by an action of G on F_0 , hence this additional structure is independent of the local bundle chart.
 2. Morphisms between bundles with the same structure group are also defined so as to “respect” this structure. Thus the smaller the G , the more structure the morphisms of this category preserve. We will use this rough idea when introducing vector bundles, orientations, and complex bundles. For example, vector bundles will have group structure $\mathrm{GL}(V)$ for some vector space V , and, as we will see, the corresponding fiber G -structure is the structure of vector space.
 3. There is one extremely special class of G -bundles that will be distinguished by the fact that their typical fiber supports *two* commuting actions of G , a left one and a right one (because the typical fiber will be G itself). After using one of them up to define the bundle, we will be able to use the above procedure to lift the other one to a global action of G on E . These bundles are called *principal*. In Section 13.7 we will show that principal G -bundles can be defined without reference to local trivializations, and in Section 14.1 that will lead us to a chart-free definition of *all* fiber bundles.

Example 7.2.6. A covering space $E \xrightarrow{\pi} M$ is a topological fiber bundle with a discrete typical fiber F . Its structure group, most generally, is the group $G = \mathrm{Diff}(F) = \mathrm{Bij}(F)$ of all permutations of points of the fiber. As we know, there is a natural $\pi_1(M, x_0)$ -action on $F = \pi^{-1}(x_0)$, which produces a homomorphism $\lambda : \pi_1(M, x_0) \rightarrow G$. Note that there is no natural action of $\pi_1(M, x_0)$ on *other* fibers of E because there is no natural way to choose the isomorphisms between the groups $\pi_1(M, x_0)$ for different choices of the basepoint x_0 . The image of λ is a subgroup of G . Since the covering space can be fully reconstructed from this

action, $\lambda(\pi_1(M, x_0))$ can be made the structure group of this bundle.

One case where there is a global action of the structure group on the whole covering space is when it is a G -principal covering space. In this case the fiber is G itself and supports a second (right) action of G which survives the fiber bundle construction.

7.3 Fiber morphisms

Now we examine morphisms (and in particular isomorphisms) between bundles with different structures on them. General bundle morphisms are allowed to be arbitrary smooth maps when restricted to a single fiber, so we will need to find a special subclass of bundle morphisms that “respects” given structures in some way. Let us first look at the simpler case of a single fiber. Say we have a G -fiber $(F, G \curvearrowright F_0, \mathcal{G}_0)$ and a G' -fiber $(F', G' \curvearrowright F'_0, \mathcal{G}'_0)$. Then any map $h : F \rightarrow F'$ can be represented by a map $\hat{h} : F_0 \rightarrow F'_0$ by using any two diffeomorphisms $\chi \in \mathcal{G}_0 \subset \text{Diff}(F, F_0)$ and $\chi' \in \mathcal{G}'_0 \subset \text{Diff}(F', F'_0)$ as follows:

$$\hat{h} = \chi' \circ h \circ \chi^{-1}. \quad (7.3.1)$$

Different choices of χ and χ' are equivalent to replacing \hat{h} in this formula with $g' \circ \hat{h} \circ g^{-1}$ for some $(g, g') \in G \times G'$. Therefore we notice the importance of the action of the product group $G \times G'$ on the set $C^\infty(F_0, F'_0)$ by compositions:

$$G \times G' \curvearrowright C^\infty(F_0, F'_0) : ((g, g'), \hat{h}) \mapsto g' \circ \hat{h} \circ g^{-1}. \quad (7.3.2)$$

A choice of a class of morphisms \tilde{h} then corresponds to a choice of a set of orbits of this action. For this we either need F_0 and F'_0 to carry additional structure (vector spaces, metric spaces, etc.) that provides such a choice in a canonical way, or to just specify the choice manually. If want to make these choices consistently across all possible pairs of model spaces (F_0, F'_0) , this amounts to forming a category. We detail these methods below:

1. Let \mathcal{S} be a subcategory of the category Man^∞ of smooth manifolds (it need not include neither all objects nor all morphisms). Then if the model spaces F_0 and F'_0 are objects of \mathcal{S} , we denote

$$\mathcal{S}(F_0, F'_0) = \text{Mor}_{\mathcal{S}}(F_0, F'_0). \quad (7.3.3)$$

We think of the category \mathcal{S} as describing “additional structure” on the model spaces. We now define morphisms from $(F, G \curvearrowright F_0, \mathcal{G}_0)$ to $(F', G' \curvearrowright F'_0, \mathcal{G}'_0)$ as maps $h : F \rightarrow F'$ which, under some (and therefore any) diffeomorphisms $\chi \in \mathcal{G}_0$ and $\chi' \in \mathcal{G}'_0$, are represented by maps $\hat{h} : F_0 \rightarrow F'_0$ belonging to $(G \times G') \cdot \mathcal{S}(F_0, F'_0)$.

2. In the situation of item 1, it is also often the case that G and G' act by automorphisms, i.e. by elements of $\text{Aut}_{\mathcal{S}}(F_0)$ and $\text{Aut}_{\mathcal{S}}(F'_0)$, respectively. Then

$$(G \times G') \cdot \mathcal{S}(F_0, F'_0) = \mathcal{S}(F_0, F'_0). \quad (7.3.4)$$

This situation comes with an extra benefit. Since G preserves the additional structure on F_0 , this structure can be “lifted” via any $\chi \in \mathcal{G}_0$ so that F itself naturally becomes an object of \mathcal{S} . Thus, in addition to the action of $Z(G)$, the fiber will inherit some additional structure from the G -invariant structure on the model fiber. We will soon see how this works explicitly on the example of vector bundles.

3. The most common setting is a special sub-case of 2, whereby G and G' are *defined* as the groups of automorphisms of F_0 and F'_0 , respectively, so we don't even need to specify them separately:

$$G = \text{Aut}_{\mathcal{S}}(F_0), \quad G' = \text{Aut}_{\mathcal{S}}(F'_0). \quad (7.3.5)$$

The actions of G and G' in this case are given canonically, and we can proceed as in 1. In particular, this implies that every representative \hat{h} of every *automorphism* (invertible morphism of a G -fiber to itself) h is an element of G .

Definition 7.3.1 (Category of \mathcal{S} -fibers). Given a G -fiber $(F, G \curvearrowright F_0, \mathcal{G}_0)$, a G' -fiber $(F', G' \curvearrowright F'_0, \mathcal{G}'_0)$, and a subset $\mathcal{S}(F_0, F'_0) \subset C^\infty(F_0, F'_0)$ determined from a category \mathcal{S} in one of the ways described above, a morphism between them is a smooth map $h : F \rightarrow F'$ such that its representative \hat{h} defined by (7.3.1) using any $\chi \in \mathcal{G}_0$ and $\chi' \in \mathcal{G}'_0$ belongs to the orbit of an element of $\mathcal{S}(F_0, F'_0)$ under the action (7.3.2) of $G \times G'$ on $C^\infty(F_0, F'_0)$. We say that h is *modeled on* an element of $\mathcal{S}(F, F')$. We call this category $\mathcal{S}\text{-Man}^\infty$ and its objects \mathcal{S} -fibers. Note that \mathcal{S} is not enough to define this category – one also needs to specify groups G and their actions (unless $G = \text{Aut}(F_0)$ as in case 3 above).

In particular, two G -fibers $(F, G \curvearrowright F_0, \mathcal{G}_0)$ and $(F', G \curvearrowright F_0, \mathcal{G}'_0)$ with the same structure group G and model space F_0 are isomorphic if and only if there exists a diffeomorphism $h : F \rightarrow F'$ such that the map $\mathcal{G}'_0 \rightarrow \mathcal{G}_0$ given by the composition $\chi' \mapsto h^*\chi' = \chi' \circ h$ is a bijection. If $G = \text{Aut}_{\mathcal{S}}(F_0)$, then this is equivalent to $\hat{h} \in G$ for all representatives (7.3.1) of h .

Note that the choice of the category \mathcal{S} for a given type of structure is not unique. In fact we mostly need this definition to produce a unique notion of *isomorphism* of \mathcal{S} -fibers and later \mathcal{S} -bundles, whereas we won't really care what the rest of the morphisms look like. We now consider a number of important examples of \mathcal{S} -fiber structures.

Example 7.3.2 (Linear G -structures). 1. If $\mathcal{S} = \text{FinVect}_{\mathbb{K}}$ (finite-dimensional \mathbb{K} -vector spaces), then we get the category of $\text{FinVect}_{\mathbb{K}}$ -fibers with linear maps. Here, if the model space is V , then we define $G = \text{Aut}_{\mathcal{S}}(V) = \text{GL}(V)$. A G -structure \mathcal{G}_0 on F is an equivalence class of diffeomorphisms $F \rightarrow V$ that differ by an automorphism of V . Since V is a vector space, we can use any of these diffeomorphisms $\chi \in \mathcal{G}_0$ to induce a vector space structure on F itself: for $\lambda_1, \lambda_2 \in \mathbb{K}$ and $e_1, e_2 \in F$ we define the linear

combination

$$\lambda_1 e_1 + \lambda_2 e_2 := \chi^{-1}(\lambda_1 \chi(e_1) + \lambda_2 \chi(e_2)). \quad (7.3.6)$$

This is well-defined because different χ 's differ by a linear automorphism of V . In particular, this defines an action of scalar matrices λI , $\lambda \in \mathbb{K}$, on F , which coincides with the induced natural action of the center $Z(G)$. Note that $GL(V)$ has no natural action on F because there is no natural isomorphism between F and V ! Thus the category $\mathcal{S}\text{-Fib}$ is isomorphic to \mathcal{S} itself.

In fact, one can replace the category $\text{FinVect}_{\mathbb{K}}$ with its skeleton $\text{Sk}(\text{FinVect}_{\mathbb{K}})$, whose objects are *standard* spaces \mathbb{K}^n , $n = 0, 1, \dots$, and the only nontrivial morphisms are inclusions-as or projections-onto the first n components. Indeed, every linear map between vector spaces is represented by one of these standard maps in some appropriately chosen bases. In this case an element of \mathcal{G}_0 , i.e. a *frame*, determines a basis on F and outputs the components of vectors in that basis.

2. In general, if structure groups G are defined as the automorphism groups $\text{Aut}(F_0)$, then the resulting category $\mathcal{S}\text{-Fib}$ is equivalent to \mathcal{S} itself. This allows one to define many classical structures on vector spaces in terms of a countable set of “standard” ones. This shows the importance of allowing G to not necessarily act by automorphisms. We proceed in this manner in the next several examples.
3. If $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, then we can let \mathcal{S} be the category of finite-dimensional vector spaces with a \mathbb{K} -valued dot product, or its skeleton consisting of spaces \mathbb{K}^n with the standard dot product $\sum_{i=1}^n \bar{x}_i y_i$. For each model space \mathbb{K}^n we also take its isometry group G , which is O_n , U_n , or Sp_n , respectively. Then \mathcal{S} -fibers are vector spaces F with a G -structure which allows us to lift the dot product onto F . There is a bijective correspondence between G -structures and dot products on F , so $\mathcal{S}\text{-Fib}$ is exactly the category of vector spaces with a dot product and all linear maps. G -frames in this context are just orthonormal bases. We can also replace \mathcal{S} with the category of vector spaces V with a dot product and linear morphisms whose operator norm is at most 1 (this forces isomorphisms to be isometries), and letting $G = \text{Aut}_{\mathcal{S}}(V)$. Then $\mathcal{S}\text{-Fib}$ has the same objects but fewer morphisms.
4. If we let $\mathcal{S} = \text{FinVect}_{\mathbb{K}}$ but redefine the structure group G of a model space V as $G = GL^+(V)$, i.e. orientation-preserving automorphisms, then \mathcal{S} -fibers are oriented vector spaces with linear maps between them.
5. If \mathcal{S} is again the category of standard real vector spaces \mathbb{R}^n , $n = 0, 1, \dots$, and the structure group of \mathbb{R}^n is $SL_n(\mathbb{R})$, then \mathcal{S} -fibers are vector spaces with a volume form (a totally antisymmetric multilinear form of maximum rank) with linear maps between them. Alternatively we can let \mathcal{S} consist of \mathbb{R}^n 's with the standard volume form v_n and linear maps $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ such that $f^* v_m = \lambda v_n$ with $\lambda \in (-1, 1]$ (this forces automorphisms to have determinant 1), and define $G = \text{Aut}_{\mathcal{S}}(\mathbb{R}^n) \cong SL_n(\mathbb{R})$. Note that

- a volume form induces an orientation because $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n^+(\mathbb{R})$.
6. Taking standard model spaces \mathbb{K}^n with G defined as the group that leaves the subspace \mathbb{K}^p spanned by the first p standard basis vectors, and linear maps that map the first p directions of one space to the first p directions of the other (not necessarily surjectively), we get the category $\mathcal{S}\text{-Fib}$ of vector spaces with “ p -directions”. It is isomorphic to the category of pairs (V, W) where W is a p -dimensional subspace of V , and morphisms are required to be maps of pairs (i.e. W is mapped inside W'). Note that morphisms between standard spaces $\mathbb{R}^n \rightarrow \mathbb{R}^m$ are represented by matrices with the top right $p \times (n - p)$ block filled with zeroes (unless $p > m$, in which case it's just the right $m \times (n - p)$ block).
 7. Consider all free actions of the group $G = \mathrm{GL}_n(\mathbb{Z})$ on n -dimensional vector spaces V . This means that G is the group of linear automorphisms of V that preserve some rank- n lattice Λ spanning V (i.e. $\Lambda = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n$ for some basis v_1, \dots, v_n). Then $\mathcal{S}\text{-Fib}$ is the category of vector spaces with an *integral affine structure*. It is isomorphic to the category of pairs (V, Λ) and linear maps mapping lattice to lattice.
 8. \mathcal{S} be the category of even-dimensional real vector spaces V with a linear isomorphism $J : V \rightarrow V$ such that $J^2 = -\mathrm{id}_V$ and linear maps $h : (V, J) \rightarrow (V', J')$ such that $h \circ J = J' \circ h$. Each J defines a complex structure on V by $iv := J(v)$. Then $G = \mathrm{Aut}_{\mathcal{S}}(V)$ consists of matrices $A \in \mathrm{GL}(V)$ such that $JA = AJ$ and is isomorphic to $\mathrm{GL}_n(\mathbb{C})$ (exercise). $\mathcal{S}\text{-Fib}$ is isomorphic to \mathcal{S} itself and is the category of complex structures on real vector spaces with complex-linear maps between them. \mathcal{S} can also be replaced with its skeleton, which consists of standard complex structures on \mathbb{R}^{2n} :

$$J_{2n} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \quad (7.3.7)$$

And linear maps $f : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2m}$ that satisfy $J_{2m} \circ f = f \circ J_{2n}$. In the case $m = n$ such maps are called holomorphic and their matrices have the form

$$M = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}, \quad (7.3.8)$$

which can be identified with $A + iB \in \mathrm{Mat}_n(\mathbb{C})$. Note that each complex structure also induces an orientation because matrices $A \in \mathrm{GL}_n(\mathbb{C}) \subset \mathrm{GL}_{2n}(\mathbb{R})$ have positive determinant over \mathbb{R} , namely $\det M = |\det(A + iB)|^2 > 0$.

9. Consider standard spaces \mathbb{R}^{2n} with standard symplectic forms, i.e. antisymmetric nondegenerate bilinear forms $\omega_{2n} : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$ whose matrix in the standard basis is J_{2n} . Let morphisms be all linear maps and let $G = \mathrm{Sp}_{2n}(\mathbb{R})$ be the group of symplectic matrices, i.e. matrices $A \in \mathrm{GL}_{2n}(\mathbb{R})$ such that $A^T J_{2n} A = J_{2n}$. Then \mathcal{S} -fibers are symplectic vector spaces (V, ω) with all linear maps between them. This can be

reduced to just the category of *symplectomorphisms*, i.e. isomorphisms that preserve the symplectic structure.

Example 7.3.3 (Hermitian structures). A Hermitian structure on a complex vector space $V_{\mathbb{C}}$ is just a symmetric sesquilinear form $\mathbf{h} : V_{\mathbb{C}} \times V_{\mathbb{C}} \rightarrow \mathbb{C}$ as introduced in example 2, i.e.

$$h(v, w) = \overline{h(w, v)}, \quad v, w \in V_{\mathbb{C}}, \quad (7.3.9)$$

and $h(v, v) \geq 0$ for all $v \in V_{\mathbb{C}}$ and vanishes iff $v = 0$.

An n -dimensional complex vector space is naturally isomorphic to a $2n$ -dimensional real vector space with a complex structure $V_{\mathbb{C}} \cong (V, J)$ (defined by $J(v) := iv$), also denoted V_J . Then \mathbf{h} can be reinterpreted as a complex-valued \mathbb{R} -bilinear form on V_J which in addition satisfies

$$\mathbf{h}(v, Jw) = -\mathbf{h}(Jv, w) = i\mathbf{h}(v, w), \quad v, w \in V_J, \quad (7.3.10)$$

and such that $\mathbf{h}(v, v)$ is a norm. This also implies that

$$\mathbf{h}(Jv, Jw) = \mathbf{h}(v, w). \quad (7.3.11)$$

We can decompose \mathbf{h} into its real and imaginary parts, which we suggestively call \mathbf{g} and ω :

$$\mathbf{h}(v, w) = \operatorname{Re} \mathbf{h}(v, w) + i \operatorname{Im} \mathbf{h}(v, w) = \mathbf{g}(v, w) + i\omega(v, w). \quad (7.3.12)$$

We have

$$\mathbf{g}(v, w) = \operatorname{Re} \mathbf{h}(v, w) = \operatorname{Re} \mathbf{h}(w, v) = \mathbf{g}(w, v), \quad (7.3.13)$$

and $\mathbf{g}(v, v) = \mathbf{h}(v, v)$, therefore \mathbf{g} is a nondegenerate symmetric bilinear form that defines a dot product on V_J . Furthermore, the sesquilinear condition implies that $\omega(v, w) = \frac{1}{2}(\mathbf{h}(w, v) - i\mathbf{h}(v, w)) = \frac{1}{2}(\mathbf{h}(w, Jv) - \mathbf{h}(v, Jw))$ is a real antisymmetric bilinear form. Moreover,

$$\omega(v, Jw) = \frac{1}{2}(\mathbf{h}(Jw, Jv) + \mathbf{h}(v, w)) = \frac{1}{2}(\mathbf{h}(w, v) + \mathbf{h}(v, w)) = \operatorname{Re} \mathbf{h}(v, w) = \mathbf{g}(v, w), \quad (7.3.14)$$

therefore ω is nondegenerate as well, and hence determines a symplectic structure on V_J . Crucially, fixing a complex structure J , the above formulas relating any two of the three objects $\mathbf{h}, \mathbf{g}, \omega$ uniquely determine the third. We thus have natural isomorphisms between the following three pairs of structures:

1. Hermitian structures (J, \mathbf{h}) on V .
2. Pairs (J, \mathbf{g}) of a complex structure J and a metric \mathbf{g} that are compatible in the sense that

$$\mathbf{g}(Jv, Jw) = \mathbf{g}(v, w), \quad \text{or} \quad \mathbf{g}(Jv, w) = -\mathbf{g}(v, Jw), \quad v, w \in V. \quad (7.3.15)$$

This compatibility condition is equivalent to the bilinear form $\omega_J(v, w) = \mathbf{g}(Jv, w)$ being a symplectic form, also called the *fundamental form*.

3. Pairs (J, ω) of a complex structure J and a symplectic structure ω that are compatible in the sense that

$$\omega(Jv, Jw) = \omega(v, w), \quad \text{or} \quad \omega(Jv, w) = -\omega(v, Jw), \quad v, w \in V, \quad (7.3.16)$$

and in addition that the resulting form $g_J(v, w) = \omega(v, Jw)$, which is symmetric because of compatibility, is positive definite (in this case we say that J *tames* ω). These conditions are equivalent to the bilinear form g_J being a real scalar product.

4. Pairs (g, ω) of a positive definite metric g and a symplectic form ω . By non-degeneracy of g , there exists a unique operator J such that $\omega(v, w) = g(v, Jw)$, and it follows that J is a complex structure compatible with both g and ω .

Thus the group of isometries of h , $U(V, h)$, coincides with the following double intersections of groups $Sp(V, \omega)$ (automorphisms preserving ω), $GL(V_{\mathbb{C}})$ (automorphisms preserving J), and $O(V, g)$ (automorphisms preserving g):

$$U(V, h) = Sp(V, \omega) \cap GL(V_{\mathbb{C}}) = O(V, g) \cap GL(V_{\mathbb{C}}) = O(V, g) \cap Sp(V, \omega). \quad (7.3.17)$$

In particular, for standard model spaces $(\mathbb{R}^{2n}, J_{2n})$ we get the “*2 out of 3 property of unitary groups*”:

$$U_n = Sp_n(\mathbb{R}) \cap GL_n(\mathbb{C}) = O_{2n} \cap GL_n(\mathbb{C}) = O_{2n} \cap Sp_n(\mathbb{R}), \quad (7.3.18)$$

which is also obvious from the fact that any two of the defining equations $A^TJA = J$, $A^{-1}JA = J$, $A^TA = I$, imply the third. Here, $GL_n(\mathbb{C})$ is included in $GL_{2n}(\mathbb{R})$ as the set of invertible matrices of the form (7.3.8) identified with $A + iB \in GL_n(\mathbb{C})$.

The triple (J, g, ω) consisting of compatible complex, metric, and symplectic structures on V can be called a *linear Kähler structure*. As we have seen, the set of Kähler structure is bijective with the set of Hermitian structures, which is why in the linear case the term Kähler structure is redundant.

- Exercise 7.3.**
1. Show that the space of metrics on a vector space is convex, i.e. for any two metrics g_1 and g_2 , all bilinear forms $tg_1 + (1-t)g_2$, $t \in [0, 1]$, are also metrics. Hence it is contractible.
 2. Show that the space of Hermitian forms h is also convex and hence contractible.
 3. Show that the set $\mathcal{J}(V, \omega) \subset GL(V)$ of complex structures compatible with a given symplectic form ω (but not necessarily taming it) is contractible. See [16, Prop. 7.5.7].

7.4 Bundle morphisms

Now we extend our definition of morphisms to arbitrary **FBS**s. Consider a morphism $\tilde{h} : E \rightarrow E'$ between two arbitrary bundles $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ with potentially differing structure groups G and G' . Pick local trivializations $\{(U_\alpha, \chi_\alpha)\}_{\alpha \in A}$, $\{(U'_\alpha, \chi'_\alpha)\}_{\alpha \in A'}$ on the

two bundles with corresponding cocycles $\{t_{\alpha\beta}\}$ and $\{t'_{ab}\}$ and denote $U_{\alpha a} = U_\alpha \cap h^{-1}(U'_a)$. Then the bundle morphism is locally represented by maps $\hat{h}_{\alpha a} : U_{\alpha a} \times F \rightarrow F'$ as follows:

$$\tilde{h} \left(\chi_\alpha^{-1}(m, f) \right) = \chi_a'^{-1} \left(h(m), \hat{h}_{\alpha a}(m, f) \right). \quad (7.4.1)$$

By transforming to other local charts χ_β and χ'_b via formula (7.2.1) we have

$$\tilde{h} \left(\chi_\beta^{-1}(m, t_{\beta\alpha}(m) \cdot f) \right) = \chi_b'^{-1} \left(h(m), t'_{ba}(h(m)) \cdot \hat{h}_{\alpha a}(m, f) \right). \quad (7.4.2)$$

On the other hand, (7.4.1) written for $\hat{h}_{\beta b}$ implies

$$\tilde{h} \left(\chi_\beta^{-1}(m, t_{\beta\alpha}(m) \cdot f) \right) = \chi_b'^{-1} \left(h(m), \hat{h}_{\beta b}(m, t_{\beta\alpha}(m) \cdot f) \right). \quad (7.4.3)$$

By comparing the right hand sides of the last two formulas and using the fact that χ'_b is a fiberwise diffeomorphism, we conclude that

$$\hat{h}_{\beta b}(m, t_{\beta\alpha}(m) \cdot f) = t'_{ba}(h(m)) \cdot \hat{h}_{\alpha a}(m, f). \quad (7.4.4)$$

This is a complicated condition on the collection of maps $\{\hat{h}_{\alpha a}\}$ with respect two different cocycles. Despite this condition these maps can still look like completely arbitrary smooth maps when restricted to a single fiber. However, in many cases we would like to consider only those bundle morphisms which, in some sense, respect the given group structures. To this end, consider the set of smooth maps $C^\infty(F, F')$ (which is an infinite-dimensional smooth manifold) and the left action of $G \times G'$ on it by compositions:

$$(G \times G') \curvearrowright C^\infty(F, F') : ((g, g'), \varphi) \mapsto g' \circ \varphi \circ g^{-1}. \quad (7.4.5)$$

Then every $\hat{h}_{\alpha a} : U_{\alpha a} \times F \rightarrow F'$ can be used to define a fiberwise $G \times G'$ -equivariant map

$$\Phi_{\alpha a} : U_{\alpha a} \times (G \times G') \rightarrow C^\infty(F, F'), \quad (7.4.6)$$

$$(m, (g, g')) \mapsto \left(f \mapsto g' \cdot \hat{h}_{\alpha a}(h(m), g^{-1} \cdot f) \right). \quad (7.4.7)$$

It is easy to see that this map is $G \times G'$ -equivariant (with $G \times G'$ acting on itself from the left). Moreover, every fiberwise $G \times G'$ -equivariant map arises this way ($\hat{h}_{\alpha a}$ can be recovered by substituting $g = g' = e$), although the correspondence of course is not bijective.²⁴ The transformation rule (7.4.4) now becomes

$$\Phi_{\beta b}(m, g, g') = \Phi_{\alpha a}(m, g t_{\beta\alpha}(m), g' t'_{ba}(h(m))). \quad (7.4.8)$$

We would like to put some restriction on the values of these maps in $C^\infty(F, F')$ that produces a well-defined class of bundle morphisms \tilde{h} which can be interpreted as “respecting” the given G - and G' -structures. This can be easily done by analogy with how we interpreted the structure of a G -fiber:

All maps $\{\Phi_{\alpha a}\}_{\alpha \in A, a \in A'}$ should be required to take values (7.4.9)

in a single orbit of the $G \times G'$ -action on $C^\infty(F, F')$. (7.4.10)

This is a well-defined constraint because, due to the equivariance of $\Phi_{\alpha a}$, the transformation law (7.4.8) implies that the fiber above p gets mapped to the same orbit regardless of the values of the indices α, a .

Definition 7.4.1 (Category $\mathcal{S}\text{-Bun}^\infty$). Let the objects of the category $\mathcal{S}\text{-Bun}^\infty$, called \mathcal{S} -bundles, be all possible structured bundles $(E \xrightarrow{\pi} M, G \curvearrowright F, \mathcal{G})$ with model fibers $F \in \text{Ob}(\mathcal{S})$ and with sets of maps $\mathcal{S}(F, F') \subset C^\infty(F, F')$ determined from a category \mathcal{S} . Morphisms from such a bundle to another one, $(E' \xrightarrow{\pi'} M', G' \curvearrowright F', \mathcal{G}')$, are bundle morphisms $\tilde{h} : E \rightarrow E'$ such that \mathcal{G} and \mathcal{G}' contain some G - and G' -atlas, respectively, in which \tilde{h} is represented by a collection of $G \times G'$ -equivariant maps $\{\Phi_{\alpha a}\}$ as above which satisfy the compatibility condition (7.4.8) and which take values in the orbit of a single element of $\mathcal{S}(F, F')$ under the action of $G \times G'$ on the set $C^\infty(F, F')$. We say that \tilde{h} is modeled on an element of $\mathcal{S}(F, F')$.

In addition, $\mathcal{S}\text{-Bun}_M^\infty$ is the subcategory of structured bundles over the fixed base manifold M with morphisms covering the identity in M , i.e. we assume $h = \text{id}_M$.

There are a number of other subcategories one may define by additionally either fixing the fiber, the structure group, or both. The most important subcategory for us will be the following.

Definition 7.4.2 (Category Bun^σ). The category Bun^σ consists of so called (G, σ) -bundles with a fixed structure group G , typical fiber F , and faithful action $\sigma : G \times F \rightarrow F$. The morphisms are required to be such that their local representatives lie in the orbit of the identity map on F . That is, locally morphisms are represented simply by actions of elements of G on F .

Example 7.4.3. Consider the cylinder $E = \mathbb{S}^1 \times \mathbb{R}$ as a line bundle over the circle. We can trivialize it with all transition functions being the identity. In the category of all line bundles (with structure group $\text{Diff}(\mathbb{R})$) there is a bundle map $E \rightarrow E$ that maps $(p, x) \mapsto (p, -x)$. However, in the category of bundles with the trivial structure group $G = \{e\}$ (i.e. the category of trivial bundles) this is not an allowed bundle map because it does not act trivially on the fiber.

Definition 7.4.4 (Bundle construction). **FBs** can be alternatively defined by specifying the base manifold M , the typical fiber F , the structure group G acting on the fiber via a *faithful*²⁵ smooth left action which can be viewed as an injective homomorphism of Lie groups $\Phi : G \rightarrow \text{Diff}(F)$ (i.e. we will write $g \cdot f = \Phi(g)(f) = \Phi_g(f)$ for $g \in G, f \in F$), a countable open covering $\{U_\alpha\}$ of M , and a cocycle $t_{\beta\alpha} \in C^\infty(U_{\alpha\beta}, G)$. Namely, we can define E as a topological space as the quotient $E := (\bigsqcup_\alpha U_\alpha \times F) / \sim$, where $(m_1, f_1) \sim (m_2, f_2)$ iff $m_1 = m_2 = m$ and $f_1 = t_{\beta\alpha}(m, f_2)$ for some α, β , and the projection π is defined as the descendent of the projections onto the first component in each summand.²⁶ Then E has a unique smooth structure that makes π a smooth submersion (see [7, Lem. 10.6] for the case of vector bundles, which is easily generalized to all **FBs**).²⁷ This construction is called the *clutching construction* of a **FB**.

The fact that any (G, σ) -bundle over M can be uniquely reconstructed from a G -valued cocycle on M means that this category doesn't really depend on the fiber F and the action σ very much. Namely, for any other faithful G -action $\sigma' : G \times F' \rightarrow F'$ there is an equivalence of categories

$$C_{\sigma', \sigma} : \text{Bun}^\sigma \rightarrow \text{Bun}^{\sigma'}. \quad (7.4.11)$$

The functors going in either direction take a bundle, choose a cocycle for it, and then construct the other bundle using the same cocycle via the other action. Faithfulness of the actions is clearly necessary for this to be an equivalence of categories because only then are the cocycles uniquely defined. The key insight flowing from this equivalence is that, to study arbitrary fiber bundles, it will suffice to pick one particular kind of typical fibers and group actions on them which makes computations easy. This computationally convenient class of bundles turns out to be the class of *principal bundles*, where the typical fiber coincides with the structure group. We discuss them in the examples below.

- Example 7.4.5** (Linear G -structures on manifolds).
1. Regardless of \mathcal{S} , if $G = \{e\}$, then a $G = \{e\}$ -structure is equivalent to a global trivialization of the bundle. In the case of vector bundles, this is called an *absolute parallelism*.
 2. $\mathcal{S} = \text{FinVect}_{\mathbb{R}}$, and the structure group of each vector space F is $\text{GL}(V)$, then \mathcal{S} -bundles are just vector bundles. Each fiber is naturally an \mathcal{S} -fiber, i.e. a vector space, and vector bundle morphisms are exactly fiberwise linear bundle morphisms.
 3. A GL_k^+ -structure on a vector bundle of rank k is called an orientation because it selects an orientation on every fiber in a continuous fashion. A vector bundle that admits such a structure is called *orientable*.
 4. An O_k -structure on a vector bundle E of rank k is equivalent to a smooth choice of metrics g_m on each fiber E_m . In Section 18.1 we will see that every real manifold admits such a choice, called a *Riemannian structure*, so every vector bundle admits a (non-unique) reduction of the structure group to an orthogonal group. O_k -bundles are

known as *Riemannian bundles*. This fact is equivalent to the statement that $\mathrm{GL}_k(\mathbb{R})$ deformation retracts onto O_n . In addition, if the bundle is orientable, the structure group can be reduced to SO_n .

5. A $\mathrm{GL}_k(\mathbb{C})$ -structure on a real vector bundle E of rank $2k$ is equivalent to a smooth choice of complex structures J_m on each fiber E_m . Such bundles are called *complex vector bundles* of rank k . Every complex vector bundle is oriented.
6. A $\mathrm{Sp}_{2k}(\mathbb{R})$ -structure on a (real) vector bundle E of rank $2k$ is equivalent to a smooth choice of symplectic forms ω_m on each fiber E_m . If such a choice exists, we get a *symplectic vector bundle*.
7. A $U_k(\mathbb{R})$ -structure on a rank k complex vector bundle is a smooth choice of a Hermitian dot product on each fiber. As we will see, every complex vector bundle admits such a choice, thus any $\mathrm{GL}_k(\mathbb{C})$ -structure can be reduced to a U_k -structure. Note that a further reduction to SU_k is not always possible, and its existence doesn't have a standard name (we will later see that it is equivalent to the triviality of the determinant line bundle $\det E$, or to the vanishing of the first Chern class, $c_1(E) = 0$).

Note that every one of these structures, just like in the linear case, defines a canonical choice of frames (bases) in each fiber, and via local trivializations they can be chosen smoothly over an open subset of the base.

Example 7.4.6 (Nonlinear G -bundles). 1. If $\mathcal{S} = \mathrm{LieGr}$ with the structure group of each Lie group G being its automorphism group $\mathrm{Aut}(G)$, then \mathcal{S} -bundles are what is called *group bundles*. Since automorphisms preserve the identity and the rest of the group structure, each fiber of a group bundle carries a natural Lie group structure. In particular, each group bundle has one canonical section $\sigma_e : M \rightarrow E$ whose value in each fiber is the identity element of that fiber.

2. If $\mathcal{S} = \mathrm{LieGr}$ but with the structure group of a Lie group G being G itself acting by left translations, then we get what are called *principal bundles* and they are so important that we shall put off a detailed discussion of them until Section 14.1. For now we only note the crucial difference between group and principal bundles: the fibers of a principal G -bundle, despite being diffeomorphic to G , are not Lie groups since left translations don't preserve the identity element. This is another important example where the group action on the typical fiber falls outside of the set of morphisms allowed by the category \mathcal{S} .

The following theorem is one of the most important theorems for the applications of fiber bundle theory, for example in complex-analytic problems of differential equations (where it is known as the Riemann-Hilbert factorization problem).

Theorem 7.4.7. *Two G -bundles E and E' over M are isomorphic iff there exists a covering $\{U_\alpha\}$ of M , two corresponding cocycles $\{t_{\alpha\beta}\}$ and $\{t'_{\alpha\beta}\}$, and a set of G -valued functions*

$\lambda_\alpha : U_\alpha \rightarrow G$ such that $t'_{\alpha\beta}(p) \equiv \lambda_\alpha(p)t_{\alpha\beta}(p)\lambda_\beta^{-1}(p)$ for all $p \in U_{\alpha\beta}$. In particular, a bundle is trivial iff its transition functions can be factorized as $t_{\alpha\beta} = \lambda_\alpha\lambda_\beta^{-1}$.

Proof. Given two bundle atlases over M , we can always take all possible pairwise intersections of the sets of the two covers of M to refine the atlases so that they are defined over the same covering of M . **WLOG**, any isomorphism of bundles over the same base can be assumed be identity on the base itself (e.g. because if the horizontal part of the isomorphism is some diffeo $f : M \rightarrow M$, we can replace one of the bundles by the isomorphic pullback bundle along f , see Definition 7.6.1). Then from (7.4.4) we see that any bundle isomorphism gives rise to the λ_α 's, and vice versa, λ_α 's completely describe the bundle isomorphism. ■

Definition 7.4.8 (Equivalent cocycles). Two G -valued cocycles on M that are related (after being brought over to a common set of charts $\{U_\alpha\}$) by the above formulas for some collection of maps $\lambda_\alpha : U_\alpha \rightarrow G$ are called *G -equivalent*. The equivalence class $[\{t_{\beta\alpha}\}]$ itself is also called a cocycle.

Corollary 7.4.9. *Given a manifold M and a faithful smooth action $G \xrightarrow{\sigma} F$, there is a bijective correspondence between isomorphism classes of (G, σ) -bundles over M and $(G$ -equivalence classes of) G -valued cocycles on M .*

This correspondence can actually be restated as an equivalence of the category Bun^σ and a certain category of cocycles.

Remark 7.4.10. Faithfulness of the action is crucial here. As we noted in Remark 7.2.2, any structure group whose action is induced from the action of $\text{Diff}(F)$ acts on F faithfully. We can see the importance of this property when reconstructing a bundle from a cocycle acting on a model fiber. If, for example, F is a single point, then every group action on F is trivial and there is only one fiber bundle over M with fiber F regardless of the structure group G , whereas there are still many non-equivalent G -valued cocycles on M . Therefore faithfulness is needed so that the cocycle is well-defined in this sense. In the future we will run into important situation where, for instance, F is a finite-dimensional vector space with an action of a non-linear group G (namely a spin group), so that this action cannot be faithful, and thus the correspondence with classes of cocycles will not be one-to-one.

7.5 Flat bundles

Now let us try to use this transformation rule for transition functions to “simplify” the cocycle and see how far we can go. If we have two charts with a transition function $t_{\alpha\beta}$, then we always find a pair of functions $\lambda_\alpha, \lambda_\beta$ on those charts such that $\lambda_\alpha t_{\alpha\beta} \lambda_\beta^{-1}$ is constant, in fact equal to identity: just pick $\lambda_\alpha = e$ and let λ_β be an arbitrary smooth extension of $t_{\alpha\beta}$ from

$U_{\alpha\beta}$ to all of U_β . We can try to keep expanding this procedure to more and more charts, but there is no guarantee that we will be able to make all transition functions constant. Thus we single out the following important class of bundles.

Definition 7.5.1 (Flat G -structure). A G -structure on a **FB** is called *flat* if it contains a bundle atlas with constant (or locally constant) transition functions. Alternatively, a G -structure is flat if structure group G can be reduced to a subgroup with a *totally disconnected* topology (one where all connected subsets are one-point sets, e.g. the discrete topology; then the transition functions are forced to be constant by continuity). If the default $\text{Diff}(F)$ -structure is flat, then the bundle itself is called flat.

Remark 7.5.2. Note that the flatness of the bundle (with respect to the default $\text{Diff}(F)$ -structure) does not imply that all other G -structures on it are flat.

Example 7.5.3 (Line bundles above \mathbb{S}^1). We can use the idea of a bundle atlas and Theorem 7.4.7 to classify all line bundles above the circle up to isomorphisms. It suffices to pick one covering of the circle, so we will pick a covering by the three circular arcs $U_1 = (\epsilon, 4\pi/3 - \epsilon)$, $U_2 = (2\pi/3 + \epsilon, 2\pi - \epsilon)$, and $U_3 = (4\pi/3 + \epsilon, 8\pi/3 - \epsilon)$. The transition functions in a line bundle take values in the group G of diffeomorphisms of \mathbb{R} . Each such diffeomorphism is a smooth function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that f' is never zero. Now, clearly G consists of two path-connected components, one of diffeomorphisms that preserve the orientation, i.e. $f'(x) > 0$ for all x , and others that reverse it, $f'(x) < 0$. We say that these diffeomorphisms have one of two possible *parities*. Let $g_+(x) = \text{id}_{\mathbb{R}}(x) = x$ and $g_-(x) = -x$, so that g_+ is the identity element of $\text{Diff}(\mathbb{R})$ and g_- is an orientation-reversing involution.

There are three transition functions, t_{12} , t_{23} , and t_{31} . Any three G -valued functions $\lambda_1, \lambda_2, \lambda_3$ provide an isomorphic line bundle. Choose $\lambda_1 \equiv g_+$. Furthermore, $t_{12} : U_{12} \rightarrow G$ can be extended to a continuous function $\lambda_2 : U_2 \rightarrow G$, $\lambda_2|_{U_{12}} = t_{12}$. Then $t'_{12} = \lambda_1 t_{12} \lambda_2^{-1} \equiv g_+$. By continuity, all values of λ_2 have the same parity. Now, we let λ_3 be a continuous extension to U_3 of the function $\lambda_2 t_{23} : U_{23} \rightarrow G$. Then $t'_{23} = \lambda_2 t_{23} \lambda_3^{-1} \equiv g_+$. Finally, $t'_{31} = \lambda_3 t_{31} \lambda_1 = \lambda_3 t_{31}$. The parity of the values of t'_{31} by construction is the product of the parities of t_{12} , t_{23} , and t_{31} .

The values of λ_3 are already uniquely fixed up to the point $2\pi - \epsilon = -\epsilon$. Due to the path-connectedness of the two components of G , λ_3 can be deformed on $U_3 \setminus U_{23} = (-\epsilon, 2\pi/3 - \epsilon)$ so that it remains smooth and $\lambda_3|_{(-\epsilon, 2\pi/3 - \epsilon)} = g_{\pm} t_{31}^{-1}$, which leads to $t'_{31} \equiv g_{\pm}$.

This means that any line bundle above \mathbb{S}^1 is isomorphic to one of two bundles, whose transition functions are either all identities, or one of them is a reflection $x \mapsto -x$. In other words, any line bundle over \mathbb{S}^1 can be reduced to the structure group $O(1) = \mathbb{Z}_2$. The former bundle is trivial, and the latter bundle is homeomorphic to the Möbius band. There are only two isomorphism classes of line bundles above \mathbb{S}^1 .

Exercise 7.4. Show that if $f : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is a diffeomorphism such that $f(0) = 0$, then $f_t(x) = f(tx)/t$ is a homotopy from f to an element of $\text{GL}_k(\mathbb{R})$. Use this to argue that the group $\text{Diff}(\mathbb{R}^k)$ has only two path-connected components, and that there are only two isomorphism classes of **FBs** with fiber \mathbb{R}^k above \mathbb{S}^1 for any k .

Exercise 7.5. Show that every diffeomorphism of the circle $\mathbb{S}^1 \subset \mathbb{C}$ is homotopic to one of the two maps $z \mapsto z^{\pm 1}$. Using this fact, classify all circle bundles over \mathbb{S}^1 (i.e. with typical fiber \mathbb{S}^1). Note that this result can also be obtained from Exercise 7.4.

Example 7.5.4. Staying in the category of topological fiber bundles, let us classify two-fold covering spaces of the “figure eight space” in the language of fiber bundles. The fiber F in this case consists of two points. The base space $M = \mathbb{S}^1 \vee \mathbb{S}^1$ can be covered by three contractible sets: one \times and two intervals. All of their overlaps are contractible, and therefore by adjusting the transition functions we will always be able to make them constant. This we have four overlaps with four corresponding elements of the structure group $G = S_2$. The bundle will be constructed out of the disjoint union of two copies of each of our three pieces:

we need to glue the following six pieces $\begin{array}{c} | \times | \\ | \times | \end{array}$. Now note that two of the transition functions, namely one on the left side of the base manifold ∞ and one on the right, can be turned into the identity by permuting the two pairs of straight segments to match those two permutations. Thus we end up with two copies of ∞ , where each loop has been cut at one point, giving us two pairs of cuts to glue back using the two permutations. Since each permutation can only take two values, we get a total of four covering spaces. Two of them happen to be isomorphic if we allow for nontrivial diffeomorphisms of the base (the left-right symmetry of ∞), but usually the classification is done without considering such symmetries. One of the fiber bundles obtained this way is trivial, and the other three are represented in Figure 4.5(1,2).

Example 7.5.5 (Covering spaces and flat bundles with flat connections). Covering spaces are just **FBs** with discrete fibers, therefore their structure group G is the permutation group of the points of the typical fiber. The classification theorem for covering spaces stated that all bundles over M with the discrete fiber F are classified by all possible homomorphisms $\pi_1(M) \rightarrow G$ (up to global conjugation by an element of G corresponding to a shuffling of the points of F). Therefore, up to the fact that the action of $\pi_1(M)$ is often not faithful, we can say that the structure group of any covering space over M is $\pi_1(M)$.

Continuing Remark 4.1.17, we see that the monodromy, or holonomy, is also nothing but a representation of $\pi_1(M)$ on the fiber. Therefore bundles with discrete fibers are completely classified by the monodromy/holonomy of the connection on them. Can we generalize these ideas to a larger class of bundles?

First of all, by a simple extension of the argument in Example 7.5.3, we see that all bundles with fiber F and structure group G above \mathbb{S}^1 are classified by homomorphisms

$\text{Hom}(\pi_1(M), \pi_0(G))/\sim$, where $\pi_0(G) = G/G_0$ and G_0 is the connected component of the identity in G (which is a normal subgroup), so that $\pi_0(G)$ is the group whose elements are the connected components of G . The equivalence relation is the global conjugation of all values of the given homomorphism by an arbitrary element of the target group (corresponds to a diffeomorphism on F , which of course doesn't change the bundle). For instance, for $F = \mathbb{R}$ and $G = \text{Diff}(\mathbb{R})$, we found that $\pi_0(G) = \mathbb{Z}_2$. The statement also agrees with the classification theorem for covering spaces.

Recall that a connection is just a prescription for lifting curves from the base into the total space. The lifted curves are called *horizontal*, and the lifting procedure is also known as *parallel transport* (of the starting point in the bundle along the path in the base). If the curve in the base is a loop γ such that $\gamma(0) = \gamma(1) = p \in M$, then its lifts with all possible starting points $\tilde{\gamma}(0) = e_0 \in E_p$ determine the holonomy $e_0 \mapsto e_1 = \tilde{\gamma}(1) = h(\gamma)(e_0)$, where we introduced the map $h : \gamma \mapsto h(\gamma) \in G$. Defined this way, $h(\gamma)$ is an element of the structure group G , and moreover h is a homomorphism from the group $L_p(M)$ of all loops based at p to G . The image of this homomorphism is a subgroup $\text{Hol}_p < G$ called the *holonomy group* of the bundle based at p . If h is restricted only to contractible loops, then the corresponding group is called $\text{Hol}_p < \text{Hol}_p < G$, the *local holonomy group* at p .

If $\widehat{\text{Hol}}_p = \{\text{id}\}$ for all p , the connection is called *flat*. This is equivalent to the property that the holonomy depends only on the homotopy class of the base loop. In particular, a small simply connected neighborhood $U \subset M$ of $p \in M$ can be uniquely lifted to a local section of E passing through $e_0 \in E_p$, which is of course diffeomorphic to U . That is, E can be locally fibrated by horizontal sections, each diffeomorphic to U .

Flatness of a connection is sometimes called *integrability*, and is exactly what people mean when they speak of integrable systems in physics: the phase space of an integrable system can be fibrated by hypersurfaces that are locally spanned by physical trajectories.

Finally, we notice that the definition of flat FBs given above implies the existence of a flat connection. Namely, a loop γ in M that passes through charts $U_1, U_2, \dots, U_n = U_1$ with the starting point $(p, f) \in U_1 \times F$ can be lifted as $(\gamma(t), f)$ on U_1 , $(\gamma(t), \tau_{21}f)$ on U_2 , $(\gamma(t), \tau_{32}\tau_{21}f)$ on U_3 , and so on (here $\tau_{ij} \in G$ represent the constant transition functions). These actually stitch together into a continuous path under the identifications in the bundle. Similarly one can lift non-closed curves, which completes the definition of the connection. Exercise: check that the local holonomy of the connection just described is independent of the choice of a bundle atlas (even though the connection itself does depend on the atlas) and trivial, i.e. the connection is indeed flat.

The non-local part of the holonomy is described by ordered products of transition functions $\prod \tau_{ij}$ along non-contractible loops. Since these are homotopy-invariant, we have a homomorphism $\pi_1(M) \rightarrow G$ that completely describes the holonomy of the flat connection. A global conjugation $g \mapsto g_0 g g_0^{-1}$ on the group doesn't affect the connection (it is essentially

equivalent to shifting the base point in the definition of $\pi_1(M)$), therefore all flat connections on bundles with structure group G above M are classified by $\text{Hom}(\pi_1(M), G)/\sim$, where \sim is the conjugation equivalence. Moreover, one can show that every homomorphism of this kind corresponds to a bundle with a connection. This is the generalization of the classification theorem for covering spaces.

Note that a flat bundle with structure group G and a flat connection is the same as a fiber bundle whose structure group is G with the discrete topology (because in the latter situation the transition functions as well as all deformation parameters λ_α are forced to be constant functions, the flat connection is unique and given by the construction above). All this leads us to conclude that flat bundles with a flat connection, or bundles with totally disconnected structure groups, are the natural generalization of covering spaces.

We have shown that bundles over M with a totally disconnected structure group G are precisely classified by the conjugacy classes of homomorphisms $\text{Hom}(\pi_1(M), G)$. The conjugacy class of such homomorphisms corresponding to a bundle $E \xrightarrow{\pi} M$ is called the *characteristic class* of the bundle. Therefore, for totally disconnected structure groups, a single characteristic class determines the isomorphism class of the bundle.

The main goal for Parts IV and VI of these lectures will be to develop the machinery necessary for the full classification of **FBs** with arbitrary topological structure groups, in particular by extending the theory of characteristic classes. The characteristic class we have just constructed (acting on one-dimensional cycles, i.e. closed loops) is only the first one in an infinite list of cohomology classes corresponding to a given bundle.

7.6 Subbundles, pullbacks

We've seen that bundles are, in categorical terms, pushouts (i.e. quotients of disjoint unions) of a family of trivial bundles along the trivialization maps. The categorical notion of pullback introduced in Definition 1.6.3 is also very useful in the theory of **FBs**.

Definition 7.6.1 (Pullback bundle). Let $E \xrightarrow{\pi} M$ be a smooth **FB** and $N \xrightarrow{f} M$ a smooth map. We define the *pullback of E along f* , or the bundle *induced from E by f* , denoted f^*E , in three ways:

- $f^*E = \{(n, p) \in N \times E \mid f(n) = \pi(p)\} \subset E \times N$ with the subspace topology induced from $E \times N$. Obviously we have two canonical projections onto the two components of the product, $f^*E \xrightarrow{\tilde{f}} E$ and $f^*E \xrightarrow{\pi'} N$. The latter one is the new bundle projection, equal to the restriction of the natural projection pr_1 on $N \times E$. We show below that there is a unique smooth structure on f^*E that makes π' into a smooth submersion.
- If the typical fiber of E is F and we know a set of transition functions $t_{\alpha\beta} : U_{\alpha\beta} \rightarrow G$

defining E , then f^*E can be defined as the bundle with base N , same fiber F , and transition functions $f^*t_{\alpha\beta} : f^{-1}(U_{\alpha\beta}) \rightarrow G$, where $f^*t := t \circ f$ is the definition of the pullback of a function t on M via f .

- In categorical terms, it is literally the pullback $N \times_M E$ that uses the maps $E \xrightarrow{\pi} M$ and $N \xrightarrow{f} M$ in its Cartesian square. This object indeed exists because it can be explicitly constructed using either of the next two alternative definitions. The pullback naturally comes with two canonical projections $f^*E \xrightarrow{\tilde{f}} E$ and $f^*E \xrightarrow{\pi'} N$.

$$\begin{array}{ccc} f^*E & \xrightarrow{\tilde{f}} & E \\ \pi' \downarrow & \curvearrowleft & \downarrow \pi \\ N & \xrightarrow{f} & M \end{array} \quad (7.6.1)$$

Lemma 7.6.2 ([16, Prop. 1.7.3, Rem. 1.7.4]). *Let N be a smooth manifold and $M \subset N$ a subset with the relative topology induced from N . Consider the following condition.*

(E) (Ea) $M \subset \bigcup_i V_i$,

(Eb) *for every i , $V_i \cap M$ admits a smooth structure of dimension l with respect to which it is an embedded submanifold of N .*

If condition (E) holds for some $l \in \mathbb{N}$, the smooth structures on $V_i \cap M$ induce a smooth structure on M , with respect to which M is an embedded submanifold of N . Conversely, if there exists a smooth structure on M which makes it an embedded submanifold of dimension l , then (E) holds for this l and the induced smooth structure on M coincides with the original one.

Proof. Pick local slice charts (V_i, ρ_i) on N such that $\rho(V_i \cap M) = \rho_i(V_i) \cap (\mathbb{R}^l \times \{0\})$. As a topological subspace of N , M is Hausdorff and second countable. For every i , $V_i \cap M$ is open in the relative topology. Being the restriction of a homeomorphism onto its image, $\rho_i|_{V_i \cap M}$ is a homeomorphism onto its image. Since by assumption the image is $\rho_i(V_i) \cap (\mathbb{R}^l \times \{0\})$ and is hence open in $\mathbb{R}^l \times \{0\}$, these data give a local chart on M . The transition maps are obtained by restriction of the original local charts on N to subsets of the subspace $\mathbb{R}^l \times \{0\}$ which are open by condition (E2). Hence the transition maps are smooth and define a smooth structure. By definition of the relative topology the inclusion is open onto its image, thus M is embedded. ■

Proposition 7.6.3 ([16, Prop. 2.6.1]). *1. f^*E defined via any of the first two definitions has a unique smooth structure that turns the bundle projection into a smooth submersion.
2. All three definitions are isomorphic as smooth bundles.*

3. f^*E , defined as in the first definition, is in fact an embedded submanifold of $N \times E$.

Proof. We apply Lemma 7.6.2. Choose a typical fiber F and a system of local trivializations $\{(U_\alpha, \chi_\alpha)\}$ for E . For every α consider the open subset $V_\alpha = f^{-1}(U_\alpha)$ of N and the mapping

$$\psi_\alpha : V_\alpha \times F \rightarrow N \times E, \quad \psi_\alpha(m, u) = (m, \chi_\alpha^{-1}(f(m), u)). \quad (7.6.2)$$

Since ψ_α is obtained by composing the diffeomorphism χ_α^{-1} with the natural inclusion mapping of the graph of $f|_{V_\alpha} : V_\alpha \rightarrow U_\alpha$, it is a smooth embedding (see Example 3.9.2). Hence the image $\psi_\alpha(V_\alpha \times F)$ inherits a smooth structure from $V_\alpha \times F$ with respect to which it is an embedded submanifold of $N \times E$. Since the image is $f^*E \cap (V_\alpha \times \pi^{-1}(U_\alpha))$ and since the $V_\alpha \times \pi^{-1}(U_\alpha)$ are open subsets of $N \times E$ covering f^*E , we conclude that f^*E is an embedded submanifold. This proves assertion 3, and Lemma 7.6.2 also shows the uniqueness of this smooth structure, proving assertion 1. Assertion 2 is an exercise. ■

Exercise 7.6 (Pullback of a bundle by itself). Given the n -fold covering of the circle $\mathbb{S}^1 \xrightarrow{\pi} \mathbb{S}^1$, show that the pullback $\pi^*\mathbb{S}^1 = \mathbb{S}^1 \times_{\mathbb{S}^1} \mathbb{S}^1$ is a trivial fiber bundle over \mathbb{S}^1 . As we will see later, this is a general fact about *principal* fiber bundles.

Remark 7.6.4. 1. From the proof of Proposition 7.6.3, every local trivialization of E induces a local trivialization of f^*E . In particular, the pullback of a trivial bundle is trivial. As indicated above, transition functions also get simply pulled back by $f^*\tau_{\alpha\beta} = \tau_{\alpha\beta} \circ f$.
2. A bundle morphism $H : E' \rightarrow E$ covering a map $h : M' \rightarrow M$ naturally decomposes as

$$E' \xrightarrow{H_{\text{ver}}} h^*E \xrightarrow{H_{\text{hor}}} E, \quad (7.6.3)$$

where $H_{\text{ver}}(x) = (\pi'(x), H(x))$, $x \in E'$, and $H_{\text{hor}} = \tilde{h}$ is the induced vector bundle morphism \tilde{h} defined by the pullback diagram (7.6.1). One can check that H_{ver} is a vertical bundle morphism and that f^*E is an embedded submanifold of $N \times E$. Using this decomposition, one can give the following characterization of isomorphisms in terms of their projections and fiber mappings: a bundle morphism is an isomorphism iff its projection (the map in the base) is a diffeomorphism and its fiber mappings are bijective.

Definition 7.6.5 (Restriction of a bundle). If $S < M$ is a submanifold via an immersion $i : S \rightarrow M$ and $E \xrightarrow{\pi} M$ is a smooth **FB**, then we can define the restricted bundle $E|_S := i^*E$. It is a smooth **FB** by the last proposition.

Definition 7.6.6 (Spaces of sections). A smooth section of a smooth **FB** $E \xrightarrow{\pi} M$ is a smooth map $\sigma : M \rightarrow E$ such that $\pi \circ \sigma = \text{id}_M$. The space of all sections, with the compact-open

topology, is denoted $\Gamma_M^\infty(E)$. If $U \subset M$ is an open subset, or more generally an immersed submanifold, then a *local section* over U is a section of the restricted bundle $E|_U$, and the space of all local sections over U is denoted $\Gamma_U^\infty(E) = \Gamma^\infty(E|_U)$.

Remark 7.6.7. It is not in general true that every local section can be realized as a restriction of a global section. For example, the two-fold cover of the circle $\mathbb{S}^1 \rightarrow \mathbb{S}^1$ has no global sections at all. More broadly, the properties of bundles that prevent an extension of local structures to global ones are called *obstructions*. The obstructions to the existence of global sections turn out to be of purely algebraic-topological nature and are described by cohomology, as we will see in Part VI.

Exercise 7.7.

1. Show that if E is trivial, then f^*E is trivial for any f . In other words, pullbacks of bundles are always “at least as trivial” as the original bundles.
2. Show that if (\tilde{f}, f) is a bundle map from E to E' that is also a fiberwise diffeomorphism, then $E \cong f^*E'$.
3. Show that if $S < M$ is immersed, weakly embedded, or embedded, then $E|_S$ is respectively an immersed, weakly embedded, or embedded, subbundle of E . See [16, Example 2.7.3].

Proposition 7.6.8. Let $E \xrightarrow{\pi} M$ be a smooth FB and let $\sigma : M \rightarrow E$ be a section (i.e. a smooth map such that $\pi \circ \sigma = \text{id}_M$). Then (M, σ) is an embedded submanifold of E .

Proof. It suffices to show that there exists an open neighborhood V of $\sigma(m)$ in E such that $(\sigma^{-1}(V), \sigma|_{\sigma^{-1}(V)})$ is an embedded submanifold. Since local trivializations are diffeomorphisms, this follows from the fact that the graph of any smooth map $\psi : M \rightarrow N$ is an embedded submanifold of $M \times N$. ■

Definition 7.6.9 (Fiber subbundle). A (immersed, weakly embedded, or embedded) subbundle of a smooth FB $E \xrightarrow{\pi} M$ is a FB $E' \xrightarrow{\pi'} M'$ with an injective FB morphism $\tilde{h} : E' \rightarrow E$ which turns E' into an immersed, weakly embedded, or embedded, respectively, submanifold of E . If $M = M'$ then we usually assume that \tilde{h} covers the identity in the base, and we call such subbundles *vertical*, or *subbundles over M* .

Proposition 7.6.10 ([16, Prop. 2.7.4]). Consider a bundle morphism $H : E' \rightarrow E$ covering a map $h : M' \rightarrow M$. TFAE:

1. (E', H) is a subbundle (resp. immersed, weakly embedded, embedded) of E .
2. (M', h) is a submanifold (resp. immersed, weakly embedded, embedded) of M and the fiber maps $H_m : E'_m \rightarrow E_m$ are injective for all $m \in M'$.
3. In the decomposition (7.6.3), (E', H_{ver}) is a vertical subbundle of h^*E and (h^*E, H_{hor}) is a subbundle (resp. immersed, weakly embedded, embedded) of E .

Proof. $1 \Rightarrow 2$ The fiber maps H_m are obviously injective. Since all bundles have local sections, let $\sigma' : U' \rightarrow E'$ be such a local section for E' . Then $H \circ \sigma'$ is a smooth local section over the set $h(U')$ (which may not be open) for E . Every local section over an open set is an embedding, in particular σ' , and thus its composition with another submanifold map, $H \circ \sigma'$, defines a submanifold of E of the same type as H . This submanifold is the image of a section of π , and π restricted to any section is a diffeomorphism onto the domain of that section. Thus the composition $\pi \circ H \circ \sigma' : U' \rightarrow h(U')$ is also a submanifold of the same type. Finally, since π is locally trivial, this means that every point $m \in M$ has a neighborhood U such that $h(M') \cap U$ coincides with a submanifold of the form we described above. This implies that the entire image $h(M)$ is a submanifold.

$2 \Rightarrow 3$ Since the H_m are injective, H_{ver} is injective, hence the assertion on (E', H_{ver}) holds since every vertical subbundle is embedded. The assertion on (h^*E, H_{hor}) was proven in Exercise 7.7.

$3 \Rightarrow 1$ Since vertical subbundles are embedded, this follows from transitivity of submanifold properties as earlier in this proof. ■

7.7 Homotopy lifting property

Theorem 7.7.1 (Bundles over $M \times I$ [17, Thm. 3.3.1]). *Every FB E over $M \times I$, where $I = [0, 1]$, is isomorphic to $E_0 \times I$ where $E_0 = E|_{M \times \{0\}}$. The isomorphism can be chosen so that its restriction to E_0 is the inclusion $E_0 \hookrightarrow E \times I$ as $p_0 \mapsto (p_0, 0)$.*

Proof. First we show that E can be trivialized over a locally finite open covering of the form $\{U_i \times I\}_i$.

By local triviality, for a fixed $x \in M$ and every $t \in I$, there is an open neighborhood $V_t \subset M$ of x and an open interval $t \in I_t \subset I$ such that E is trivial over $V_t \times I_t$. Therefore we can construct a countable, locally finite open covering of $M \times I$ with sets of this kind. By compactness, there is a finite set $0 < t_1 < \dots < t_k < 1$ such that $\bigcup_i I_{t_i} = I$. Let $V_i = V_{t_i}$, $I_i = I_{t_i}$, and define

$$U_i = \bigcap_{j=1}^i V_j, \quad J_i = \bigcup_{j=1}^i I_j. \quad (7.7.1)$$

Obviously E is trivial over $U_1 \times J_1$. Let χ_1 be a trivialization of E over $U_1 \times J_1$ and $\tilde{\chi}_2$ – one over $V_2 \times I_2$. Then $(U_1 \times J_1) \cap (V_2 \times I_2) = U_2 \times (J_1 \cap I_2)$. Assuming that I_2 intersects J_1 non-trivially and one is not contained in the other (otherwise I_2 can be dropped from the covering of I), consider the transition function $\tau : U_2 \times (J_1 \cap I_2) \rightarrow G$ defined by $\chi_1(p) = (\text{id} \times \tau(\pi(p))) \cdot \tilde{\chi}_2(p)$. Choose $c \in J_1 \cap I_2$ and a continuous function

$f : I_2 \rightarrow J_1 \cap I_2$ such that $f(t) = t$ for $t \geq c$ to define

$$\tilde{\tau} : U_2 \times I_2 \rightarrow G, \quad \tilde{\tau}(x, t) = \tau(x, f(t)). \quad (7.7.2)$$

Clearly τ and $\tilde{\tau}$ coincide on $U_2 \times ([0, c] \cap I_2)$. Hence the new mapping

$$\chi_2(p) = \begin{cases} \chi_1(p), & \pi(p) \in U_2 \times [0, c], \\ (\text{id} \times \tilde{\tau}(\pi(p))) \cdot \tilde{\chi}_2(p), & \pi(p) \in U_2 \times I_2 \end{cases} \quad (7.7.3)$$

defines a trivialization of E over $U_2 \times I_2$. By induction, we find that E is trivial over every $U_i \times I$.

We sketch the rest of the proof. Since this construction applies to every $x \in M$, we can extract a locally finite open covering of the entire M of the form $\{U_i \times I\}_i$ such that E is trivial over every set in the covering (this can be done using a **POU**). Let χ_i be the corresponding local trivializations and $\hat{\chi}_i$ the induced local trivializations of $E_0|_{U_i} \times I$.

Now we construct a global trivialization inductively in i . For $V_1 = U_1$, there is already an isomorphism $\Phi_1 : E|_{V_1 \times I} \rightarrow E_0|_{V_1} \times I$. We can construct a chain of open sets $V_1 \subset V_2 \subset \dots$ that eventually covers all of M and $U_i \subset V_i$, and by induction extend the isomorphism to $\Phi_i : E|_{V_i \times I} \rightarrow E_0|_{V_i} \times I$. Via the local trivializations χ_i and $\hat{\chi}_i$, bundle isomorphisms are represented by G -valued functions $g_i : U_i \times I \rightarrow G$ with $g_i(x, 0) = e \in G$. Let $g_1 : (V_1 \cap U_2) \times I \rightarrow G$ be the representative of Φ_1 in χ_2 and $\hat{\chi}_2$. Then g_1 can be smoothly extended to $V_2 \times I$ so that it equals e on $(U_2 \setminus V_1) \times I$. This represents the isomorphism Φ_2 that agrees with Φ_1 on the common domain. Repeating this procedure on U_3 and so on, we construct a global isomorphism of the bundles in question. See [17, Thm. 3.3.1] for details. ■

Corollary 7.7.2. *Continuous deformations (homotopies) of transition functions within the class of cocycles don't change the isomorphism class of a fiber bundle.*

Therefore, in the most general and abstract sense, the “full” characteristic class of a fiber bundle (which uniquely classifies the bundle) is the “conjugacy” and homotopy equivalence class of cocycles for it.

Now we are ready to prove the smooth version of the **HLP** for fiber bundles (a continuous version of it holds automatically since every smooth fiber bundle is a Hurewicz fibration).

Theorem 7.7.3 (Homotopy Lifting Property/Covering Homotopy Theorem). *Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be smooth **FBS**s with the same typical fiber F . If (\tilde{H}_0, H_0) is a bundle map from E to E' and $H : M \times I \rightarrow M'$ is a homotopy of H_0 , i.e. $H_0 = H|_{M \times \{0\}}$, then there*

exists a bundle map \tilde{H} that covers H :

$$\begin{array}{ccc} E \times I & \xrightarrow{\tilde{H}} & E' \\ \pi \times \text{id} \downarrow & \curvearrowright & \downarrow \pi' \\ M \times I & \xrightarrow{H} & M' \end{array} \quad (7.7.4)$$

Proof. Consider the map

$$\lambda : E \rightarrow H_0^* E' \subset M \times E', \quad \lambda(p) = (\pi(p), \tilde{H}_0(p)). \quad (7.7.5)$$

This is a bundle isomorphism of two bundles over M . Moreover, by Theorem 7.7.1, there exists a bundle isomorphism

$$\Phi : H^* E' \rightarrow \tilde{H}_0^* E' \times I \quad (7.7.6)$$

of bundles over $M \times I$ satisfying $\Phi((x, 0), f) = ((x, f), 0)$. Together with the natural bundle morphism $\text{pr}_2 : H^* E' \rightarrow E'$, the isomorphisms λ and Φ combine to a morphism

$$\tilde{H} : E \times I \xrightarrow{\lambda \times \text{id}_I} H_0^* E' \times I \xrightarrow{\Phi^{-1}} H^* E' \xrightarrow{\text{pr}_2} E' \quad (7.7.7)$$

that covers H . Since

$$\tilde{H}(p, 0) = \text{pr}_2 \circ \Phi^{-1}((\pi(p), \tilde{H}_0(p)), 0) = \text{pr}_2 \circ ((\pi(p), 0), \tilde{H}_0(p)) = \tilde{H}_0(p), \quad (7.7.8)$$

\tilde{H} is indeed an extension of \tilde{H}_0 . ■

Corollary 7.7.4. 1. If $E' \xrightarrow{\pi'} M'$ is a **FB** and $f, g : M \rightarrow M'$ are two smooth maps that are homotopic to each other, then the pullback bundles $f^* E'$ and $g^* E'$ are isomorphic.
 2. All **FBs** over contractible base spaces are trivial.

Proof. 1. Let $H : M \times I \rightarrow M'$ be a homotopy from f to g and consider the bundle $P = H^* E'$ over $M \times I$. Let $P_t = P|_{X \times \{t\}}$ viewed as a bundle over M . Then $P_0 = f^* E'$ and $P_1 = g^* E'$. By the above Theorem, P is isomorphic to $P_0 \times I$. By restricting the isomorphism to the subbundle $P_1 \subset P$, we get an isomorphism from P_1 to P_0 .
 2. Contractible spaces by definition are homotopy equivalent to a point, and all bundles over a point are trivial with the trivialization given by the pullback of the bundle bundle along the contraction. ■

We've established that fiber bundles are, roughly speaking, classified by homotopy-plus-conjugacy classes of cocycles. We now present the first elementary example in which this classification can be reduced to familiar algebraic objects.

Corollary 7.7.5 (Classification of fiber bundles over spheres). *Isomorphism classes of FBs over the sphere \mathbb{S}^n , $n \geq 1$, with fiber F and a given faithful action $G \curvearrowright F$ of the structure group G are in one-to-one correspondence with the homotopy classes of maps from \mathbb{S}^{n-1} to the identity component G_0 , up to conjugation by the discrete group of connected components of G , $G/G_0 \cong \pi_0(G)$:*

$$\text{Sk}(\text{Bun}_{\mathbb{S}^n}(G \curvearrowright F)) \cong [\mathbb{S}^{n-1}, G_0]/\pi_0(G) = \pi_{n-1}(G)/\pi_0(G), \quad (7.7.9)$$

(here the group $\pi_0(G)$ acts on the set $\pi_{n-1}(G)$ by conjugation of the values, which is a well-defined action within homotopy classes). In particular, when G is connected, this set forms a group isomorphic to $\pi_{n-1}(G)$.

Proof. We have already proven this for $n = 1$ in Example 7.5.3. For $n \geq 2$, \mathbb{S}^n can be covered by two hemispheres U_\pm that overlap along a neighborhood U_0 of the equator. The equator C is diffeomorphic to \mathbb{S}^{n-1} and U_0 can be chosen so that $U_0 \cong \mathbb{S}^{n-1} \times (-1, 1)$. Then the only transition function in any given bundle $E \xrightarrow{\pi} \mathbb{S}^n$ is $\tau : U_0 \rightarrow G$. It can be chosen so that it maps a given point $x_0 \in C$ to $e \in G$. Its restriction to the equator provides the *clutching function* $\theta = \tau|_C : \mathbb{S}^{n-1} \rightarrow G_0$ whose homotopy class will be the one put into correspondence with the bundle.

To check that this mapping is consistent, we need to show that isomorphic bundles have homotopic clutching functions, up to conjugation by an element of $\pi_0(G)$. Let the two transition functions be τ_1 and τ_2 and consider an isomorphism given by $\tau_2 = \lambda_+ \tau_1 \lambda_-^{-1}$, where λ_\pm are two G -valued functions on U_\pm , respectively. Since $\tau_{1,2}(x_0) = e$, we can first consider the case when $\lambda_\pm(x_0) = e$ and therefore λ_\pm take values in the connected component of the identity $G_0 \subset G$. Then we can deform λ_\pm by a homotopy so that $\lambda_+(P_+) = \lambda_-(P_-) = e \in G$, where P_\pm are the respective poles of the sphere. Since U_\pm are both diffeomorphic to a ball $B^n \in \mathbb{R}^n$, we can use the Cartesian coordinate x on the ball to introduce the homotopy $\lambda_+(tx)\tau_1(x)\lambda_-^{-1}(tx)$ from τ_1 to τ_2 .

Another choice for λ_\pm is when $\lambda_+(x_0) = \lambda_-(x_0) = g_0$ with g_0 in another connected component of G than the identity. Then the same argument leads to clutching functions that are, up to a homotopy, conjugates of each other by g_0 . Different choices of g_0 within its connected component only change the clutching functions by a homotopy, just like above, because each connected component of G is path-connected.

Finally, the map constructed is clearly surjective because every function $\mathbb{S}^{n-1} \rightarrow G_0$ can be used to construct a fiber bundle. It is also injective because any two homotopic clutching

functions $\theta_1 \sim \theta_2 : \mathbb{S}^{n-1} \rightarrow G$ can only be restrictions of two homotopic transition functions $\tau_1 \sim \tau_2 : \mathbb{S}^{n-1} \times (-1, 1) \rightarrow G$ (all up to conjugation by $\pi_0(G)$). This is because every such $\tau(x, t)$ is homotopic to one that is independent of t (and this is because every function on $(-1, 1)$ is homotopic to a constant one). ■

The above Corollary shows the importance of the computation of homotopy groups of at least the classical Lie groups. Interestingly, many of these groups can be computed *using* certain fiber bundles. We will begin to do this in Section 13.8.

Our ultimate goal for Part VI will be to extend (7.7.9)
to arbitrary base manifolds using the theory of characteristic classes. (7.7.10)

Example 7.7.6 (Vector bundles over \mathbb{S}^1). Now we can easily solve Exercise 7.4. Consider only vector bundles of rank k above \mathbb{S}^1 . The structure group is $G = \mathrm{GL}_k(\mathbb{R})$. This group has two connected components distinguished by the sign of the determinant. Therefore $\pi_0(G) = \mathbb{Z}_2$ and there are only two isoclasses of bundles. Equivalently we can classify bundles with a two-point fiber \mathbb{Z}_2 above \mathbb{S}^1 . We have $G = \mathrm{Diff}(\mathbb{Z}_2) = \mathbb{Z}_2$ and

$$\mathrm{Bun}_{\mathbb{S}^1}(\mathbb{Z}_2) \cong \pi_0(\mathbb{Z}_2) = \mathbb{Z}_2. \quad (7.7.11)$$

Example 7.7.7 (Plane bundles over \mathbb{S}^2). Let us classify all \mathbb{R}^2 -bundles over \mathbb{S}^2 . We have by Corollary 7.7.4 that

$$\mathrm{Sk}\left(\mathrm{Bun}_{\mathbb{S}^2}\left(\mathbb{R}^2\right)\right) \cong \pi_1(\mathrm{GL}_2(\mathbb{R})) / \pi_0(\mathrm{GL}_2(\mathbb{R})) = \mathbb{Z}/\mathbb{Z}_2 \cong \mathbb{Z} \quad (7.7.12)$$

Similarly we can classify circle bundles above the 2-sphere. The structure group of circle bundles is $\mathrm{Diff}(\mathbb{S}^1)$, which is homotopy equivalent to $O_2 \cong \mathbb{S}^1 \sqcup \mathbb{S}^1$. Therefore

$$\mathrm{Sk}\left(\mathrm{Bun}_{\mathbb{S}^2}\left(\mathbb{S}^1, O_2\right)\right) \cong \pi_1(O_2) / \pi_0(O_2) \cong \mathbb{Z}/\mathbb{Z}_2 \cong \mathbb{Z} \quad (7.7.13)$$

(because \mathbb{Z}_2 acts by conjugation, which is a trivial action on the abelian group \mathbb{Z}). If we were to restrict the structure group to orientation-preserving diffeomorphisms of the circle, we would need to replace O_2 by $SO_2 = U_1$, so

$$\mathrm{Bun}_{\mathbb{S}^2}(\mathbb{S}^1, U_1) \cong \pi_1(\mathbb{S}^1) \cong \mathbb{Z}. \quad (7.7.14)$$

This implies that every plane bundle over \mathbb{S}^2 is orientable. The integer number that classifies these (oriented) plane or circle bundles over the 2-sphere is called the *Chern number* and represents yet another characteristic class of bundles.

Example 7.7.8 (Hopf bundle revisited). If we cover \mathbb{S}^2 by two disks D_{\pm} overlapping along the equator, the Hopf fibration $\mathbb{S}^3 \xrightarrow{\pi} \mathbb{S}^2$ (Example 7.1.10) restricted to each of them is trivial and diffeomorphic to a solid torus in \mathbb{R}^3 . Therefore the 3-sphere \mathbb{S}^3 can be broken up into two disjoint (aside from their boundary) solid tori that are glued to each other along the 2-torus that is their common boundary. The twists of the boundary performed in this process classify the resulting circle bundles over the 2-sphere, so here we will try to understand the case of the Hopf bundle. In terms of the coordinates on the bundle used in Example 7.1.10, we have

$$E|_{D_+} = \{|z_1| \geq |z_2|\}, \quad E|_{D_-} = \{|z_1| \leq |z_2|\}. \quad (7.7.15)$$

Note that $D_+ = \mathbb{S}^2 \setminus \{\pi(0, 1)\}$ and $D_- = \mathbb{S}^2 \setminus \{\pi(1, 0)\}$. It is conventional to define $0 := \pi(1, 0)$ and $\infty := \pi(0, 1)$. Therefore we can also call these sets $U_0 := D_+$ and $U_\infty := D_-$. The standard complex coordinate on $U_0 = D_+$ is $z := \frac{z_2}{z_1}$ and on $U_\infty = D_-$, $w := \frac{z_1}{z_2}$. The transition map is $w(z) = \frac{1}{z}$, which is complex analytic, so this atlas turns \mathbb{S}^2 into a *complex manifold*, namely the projective complex line $\mathbb{C}P^1$.

Over the equator C we have

$$E|_C = \left\{ |z_1|^2 = |z_2|^2 = \frac{1}{2} \right\}, \quad (7.7.16)$$

so the restriction of the Hopf bundle to the equator is a 2-torus with two angular coordinates φ, ψ , where φ is the standard azimuthal coordinate on the sphere and ψ is the coordinate in the fiber:

$$E|_C = \left\{ (z_1, z_2) = \frac{1}{\sqrt{2}} \left(e^{i\varphi/2+i\psi}, e^{-i\varphi/2+i\psi} \right) \right\}. \quad (7.7.17)$$

Identifying D_{\pm} with \mathbb{C} via the coordinates z and w , respectively, we construct the bundle atlas

$$\chi_0(z_1, z_2) = \left(\frac{z_2}{z_1}, \frac{z_1}{|z_1|} \right) \in \mathbb{C} \times \mathbb{S}^1, \quad 0 \leq |z_2| \leq |z_1|, \quad (7.7.18)$$

$$\chi_\infty(z_1, z_2) = \left(\frac{z_1}{z_2}, \frac{z_2}{|z_2|} \right) \in \mathbb{C} \times \mathbb{S}^1, \quad 0 \leq |z_1| \leq |z_2|. \quad (7.7.19)$$

Incidentally, limiting $|z_1|$ and $|z_2|$ to $[0, \frac{1}{2}]$, this shows once again that the 3-sphere is a union of two solid tori overlapping only at their common \mathbb{T}^2 -boundary. Note that in this atlas, a point over the equator gets mapped to a conjugate pair of points on $\partial\mathbb{D}^2$ by $\text{pr}_1 \circ \chi_{\pm}$ (this could be fixed by redefining one of the charts, but that would violate the convenient analytic

structure). These trivializations are indeed diffeomorphisms because their inverses are

$$\chi_0^{-1}(z, e^{i\psi}) = \left(\frac{e^{i\psi}}{\sqrt{1+|z|^2}}, \frac{ze^{i\psi}}{\sqrt{1+|z|^2}} \right), \quad (z, e^{i\psi}) \in \mathbb{C} \times \mathbb{S}^1, \quad (7.7.20)$$

$$\chi_\infty^{-1}(z, e^{i\psi}) = \left(\frac{ze^{i\psi}}{\sqrt{1+|z|^2}}, \frac{e^{i\psi}}{\sqrt{1+|z|^2}} \right), \quad (z, e^{i\psi}) \in \mathbb{C} \times \mathbb{S}^1. \quad (7.7.21)$$

We compute the clutching function

$$\tau_{\infty 0}(z) \cdot e^{i\psi} = \text{pr}_2 \circ \chi_\infty \circ \chi_0^{-1}(z, e^{i\psi}) = \frac{z}{|z|} \cdot e^{i\psi}, \quad (7.7.22)$$

in particular, on the equator, $\tau_{\infty 0}(e^{i\varphi}) = e^{i\varphi}$. Therefore, following a standard sign convention, the Chern number of this bundle equals -1 .²⁸

Example 7.7.9 (Plane/circle bundles over \mathbb{S}^2 continued). 1. As we saw above, every plane bundle over \mathbb{S}^2 is orientable. This can be made more general. We will see later that non-orientability of a vector bundle over a manifold M implies the existence of some nontrivial two-fold covering of M . Therefore every vector bundle over a simply connected base is orientable, since such manifolds have no non-trivial covering spaces.

2. Since the transition functions of every circle bundle over \mathbb{S}^2 can always be reduced to U_1 , and U_1 is abelian, every such bundle comes with a natural global free (and fiberwise transitive) action of U_1 . This bundle can be trivialized over two hemispheres, thus being decomposed into a gluing of two solid tori $\mathbb{D}^2 \times \mathbb{S}^1$. Recall that the lens space $L(p; 1)$ (Example 2.5.32) can similarly be decomposed into two solid tori glued along a map acting on the boundary torus \mathbb{T}^2 of one of them by $(e^{2\pi i\varphi}, e^{2\pi i\psi}) \mapsto (e^{2\pi i\varphi}, e^{2\pi i(p\varphi+\psi)})$. But this is exactly the identification that produces the circle bundle over \mathbb{S}^2 with the transition function $\tau_{-+}(e^{2\pi i\varphi}) = e^{2\pi i p\varphi}$. Thus the total space of every \mathbb{S}^1 -bundle over \mathbb{S}^2 is a lens space. Note that $p = 1$, corresponding to the Hopf bundle, gives the correct total space $L(1; 1) = L(1; 0) = \mathbb{S}^3$. Another example is $p = 2$, which gives $\mathbb{R}P^3$.

Remark 7.7.10. In the mid-20th century there was a conjecture that any two closed (compact with no boundary) manifolds of the same dimension that are homotopy equivalent must be homeomorphic. 3-dimensional lens spaces were the first counterexample: $L(p; q_1)$ and $L(p; q_2)$ are homotopy equivalent iff $q_1 q_2 \equiv \pm n^2 \pmod{p}$ for some $n \in \mathbb{N}$, but homeomorphic only if $q_1 \equiv \pm q_2^{\pm 1} \pmod{p}$.

The enhanced conjecture of Armand Borel asks the same question under the additional constraint that both manifolds are aspherical (trivial homotopy groups $\pi_{\geq 2}$). The answer to the Borel conjecture is positive in 3 dimensions thanks to Perelman's solution of the

Poincaré conjecture. It is also known to hold for Riemannian manifolds of negative curvature. It remains an open problem outside of these cases, but is expected to be false in higher dimensions.

Exercise 7.8. Show that a circle bundle over \mathbb{S}^2 has a global section only if it is trivial. In particular, the Hopf bundle has no global sections.

8 Vector Bundles

8.1 Vector bundles

Vector bundles are nothing but **FBs** whose fibers carry vector space structures such that these structures, in some sense, vary smoothly with the base point. We make this precise by putting a natural restriction on the transition functions using the language of G -structures. For this definition we allow the field \mathbb{K} be \mathbb{R}, \mathbb{C} , or \mathbb{H} , but we will only deal with real vector bundles until we get to complex geometry.

Definition 8.1.1 (Vector bundle). A **vector bundle (VB)** is a $GL(V, \mathbb{K})$ -**FB**, where V is the typical fiber and also a finite-dimensional \mathbb{K} -vector space. The dimension k of the fiber is called the *rank* of the vector bundle.

As we discussed in Section 7.2, any G -structure transplants some kind of data from the model fiber F onto all of the fibers of the bundle. In the case of vector bundles this data is the structure of a vector space. We've already seen this to be true when we discussed G -structures, but we provide an explicit proof here.

Proposition 8.1.2. *Every fiber of a vector bundle is naturally a vector space.*

Proof. Consider a fiber $E_m = \pi^{-1}(m)$, $m \in M$. By definition of the typical fiber, we have a diffeomorphism $f_\alpha = \chi_\alpha|_{E_m} : E_m \rightarrow V$, which is just the restriction of a local trivialization. We can induce a vector space structure on E_m by defining $k \cdot e_1 + e_2 := f_\alpha^{-1}(k \cdot f_\alpha(e_1) + f_\alpha(e_2))$ for $k \in \mathbb{R}$ and $e_{1,2} \in E_m$. For any other local trivialization χ_β , we have $f_\alpha = \tau_{\alpha\beta} f_\beta$ with $\tau_{\beta\alpha}$ linear on V , therefore

$$\alpha \cdot e_1 + e_2 = f_\beta^{-1} \tau_{\alpha\beta}^{-1} (k \cdot \tau_{\alpha\beta} f_\beta(e_1) + \tau_{\alpha\beta} f_\beta(e_2)) = f_\beta^{-1} (k \cdot f_\beta(e_1) + f_\beta(e_2)), \quad (8.1.1)$$

which is exactly what we would get if we induced the linear structure via f_β .

Therefore by providing a diffeomorphism between a single fiber and F we can extend the vector space structure to all fibers above the connected component of M . The resulting linear structure is independent of the choice of local trivializations. ■

The notion of morphisms between **VBs** is a special case of our definition of morphisms between **S**-bundles, namely they are bundle morphisms modeled on linear maps of vector spaces. We detail it here for clarity.

Definition 8.1.3 (Vector bundle morphism). A **VB** morphism is a **FB** morphism that is fiberwise linear, i.e. is a morphism of vector spaces when restricted to any fiber. More explicitly, if we have local trivializations $\{\chi_\alpha\}$ for $E \xrightarrow{\pi} M$ and $\{\chi'_\alpha\}$ for $E' \xrightarrow{\pi'} M'$, then a

vector bundle morphism is a **FB** morphism consisting of $\tilde{h} : E \rightarrow E'$ and $h : M \rightarrow M'$ that has the form

$$\chi'_a \circ \tilde{h} \circ \chi_\alpha^{-1}(m, v) = (h(m), \hat{h}_{\alpha a}(m) \cdot v), \quad (8.1.2)$$

where $(m, v) \in U_\alpha \times F$ and $\hat{h}_{\alpha a}(m) \in \text{Hom}(V, V')$ takes values in linear operators between the vector spaces V and V' .

With this, we have the categories VectBun^∞ of all smooth **VBs** and VectBun_M^∞ of smooth **VBs** above a given base.

Example 8.1.4. The cylinder and the Möbius band are vector bundles of rank one over \mathbb{S}^1 .

Example 8.1.5 (Holomorphic line bundles over $\mathbb{C}P^1$). For $k \in \mathbb{Z}$, define the complex line bundle $\mathcal{O}(k)$ over the Riemann sphere $\mathbb{C}P^1$ with the standard atlas consisting of (U_0, z) and (U_∞, w) (see Example 7.7.8) by the fiber \mathbb{C} , structure group $\text{GL}_1(\mathbb{C}) = \mathbb{C}^\times$, and transition function

$$\tau_{0\infty}(z) = z^k. \quad (8.1.3)$$

This bundle is called *holomorphic* since its transition functions are complex analytic as maps $U_{\alpha\beta} \rightarrow G$. Note that, since the local coordinates on U_0 and U_∞ are related by $w = \frac{1}{z}$, the opposite transition function also reads $\tau_{\infty 0}(w) = w^k$, so the integer k is uniquely defined. In Example 7.7.9 we fully classified all real plane bundles over the sphere by the same integer k , therefore we conclude that every plane bundle over \mathbb{S}^2 is isomorphic, as real **VBs**, to one of the $\mathcal{O}(k)$. The number k is called the Chern number of these bundles. Note, however, that there is no canonical isomorphism between \mathbb{S}^2 and $\mathbb{C}P^1$, so over \mathbb{S}^2 the Chern number is defined only up to a sign.

As we have seen, the Hopf bundle is the circle bundle of $\mathcal{O}(-1)$, with its structure group reduced to $U_1 \subset \mathbb{C}^\times$.

Remark 8.1.6. Pullbacks of **VBs** are **VBs**. Indeed, since pullbacks can be defined via pullbacks of transition functions, this means that the transition functions of the pullback bundle are still $\text{GL}(F)$ -valued and hence preserve the vector space structure of the fibers.

Example 8.1.7. Let $M = \mathbb{S}^1$ and let $E \rightarrow M$ be the Möbius strip as a line bundle over \mathbb{S}^1 . Realize \mathbb{S}^1 as the unit circle in \mathbb{C} and consider the n -fold covering $f_n : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ given by $f_n(z) = z^n$. Since every line bundle over \mathbb{S}^1 is isomorphic either to the cylinder $\mathbb{S}^1 \times \mathbb{R}$ or the Möbius band, we conclude

$$f_n^* E = \begin{cases} E, & n \text{ odd}, \\ \mathbb{S}^1 \times \mathbb{R}, & n \text{ even}. \end{cases} \quad (8.1.4)$$

Definition 8.1.8 (Space of sections). As always, a (global) section of a **FB** $E \xrightarrow{\pi} M$ is a smooth map $\sigma : M \rightarrow E$ such that $\pi \circ \sigma = \text{id}_M$, i.e. it picks a single point in every fiber in a

smooth fashion. In the case of vector bundles, the set $\Gamma^\infty(E)$, also denoted by $\Omega^0(M, E)$ (not to be confused with the space of all smooth maps from M to E), of all sections is a vector space. $\Gamma^\infty(E)$ is naturally a $C^\infty(M)$ -module because a section $\sigma \in \Gamma^\infty(E)$ can be multiplied by a scalar function $f \in C^\infty(M)$ point-wise as $(f \cdot \sigma)(p) = f(p) \cdot \sigma(p)$. For an open neighborhood $U \subset M$, the $C^\infty(U)$ -module of local sections over U is denoted $\Gamma_U^\infty(E)$.

- Remark 8.1.9.**
1. Unlike general **FBS**, **VBs** always have at least one global section. Let $E_0 \subset E$ be the set of all points that are neutral elements of their respective fibers as vector spaces (i.e. they are mapped to zero by the second component of any local trivialization). Furthermore, the restricted projection $\pi|_{E_0} : E_0 \rightarrow M$ is a diffeomorphism because locally it is just the first component of any local trivialization, therefore a local diffeo, and it is also clearly bijective. E_0 , and the corresponding map $\sigma_0 : M \rightarrow E_0$, are called the *zero section* of E .
 2. Even on a smooth vector bundle, not every smooth local section can be extended to a global one. Consider for example the trivial bundle $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ and its local section over $(0, +\infty)$ given by $\sigma(x) = (x, \frac{1}{x})$. It cannot be extended to all of \mathbb{R} .
 3. However, for a smooth local section $\sigma : U \rightarrow E$ and every point $p \in U$ there exists an open neighborhood $U_p \subset M$ of p such that σ can be extended to a section $\tilde{\sigma} : U \cup U_p \rightarrow E$.
 4. If $S \lhd M$ is an embedded submanifold, then any section σ of the restricted **VB** $E|_S$ can be extended to an open neighborhood of S (which may be S itself if S is open). This can be done by constructing local frames of E adapted to S at every point $p \in S$, using them to extend the section to some neighborhood U_p , and stitching these extensions together using a **POU**.
 5. If $S \lhd M$ is a *closed* embedded submanifold, then every section of $E|_S$ can be extended to a global section of E . This can be done by first extending to some open neighborhood U of S , so that $U \setminus S$ is open, and then we can use **POU** to extend $\sigma \in \Gamma^\infty(E|_S)$ in such a way that it coincides with the zero section outside of some smaller open neighborhood $S \subset V \subset M$ and can thus be extended to all of M by letting it be equal to the zero section on $M \setminus U$.

Definition 8.1.10 (Extending functors from **FinVect** to **VectBun**). Functors acting on vector spaces can be extended to vector bundles. Namely, if we have a functor $\mathcal{F} : \text{FinVect} \rightarrow \text{FinVect}$ that can act at least on isomorphisms and such that its action on morphisms (matrices) is smooth in the standard smooth structure on matrix spaces, then we can construct a new bundle by replacing the typical fiber V with $\mathcal{F}(V)$ and the transition functions $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \text{GL}(V)$ by $\mathcal{F} \circ \tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \text{GL}(\mathcal{F}(V))$.

1. The dualization functor $V \mapsto V^*$ gives rise to the dual bundle E^* , whose transition functions are $(\tau_{\alpha\beta}^{-1})^*$ (inverse transposed as matrices).

2. Direct sums of vector spaces give rise to direct sums of vector bundles $E \oplus E'$ with transition functions $\tau_{\alpha\beta} \oplus \tau'_{\alpha\beta}$ (*Whitney sum*).
3. Similarly we have tensor products of bundles $E \otimes E'$ with transition functions $\tau_{\alpha\beta} \otimes \tau'_{\alpha\beta}$ (*Whitney product*).
4. For a pair of bundles over a given base manifold, we can construct the bundle $\text{Hom}(E_1, E_2)$, which can be identified with $E_1^* \otimes E_2$. We also have $\text{End}(E) = \text{Hom}(E, E)$.
5. All of the above are in fact functors on VectBun^∞ .

Example 8.1.11. Let E_1, E_2 be **VBs** over M and let $E_1 \times E_2$ be the product bundle over $M \times M$. Let $\Delta : M \rightarrow M \times M$ be the diagonal mapping $\Delta(m) = (m, m)$. The pullback $\Delta^*(E_1 \times E_2)$ is naturally isomorphic to the direct sum $E_1 \oplus E_2$.

Definition 8.1.12 (Frames). A local frame on a **VB** $E \xrightarrow{\pi} M$ of rank k is an open set $U \subset M$ and a collection of local sections $e_i : U \rightarrow E$, $i = 1, \dots, k$ such that for every $m \in U$ the set $\{e_i(m)\}_{i=1}^k$ is a basis of the fiber E_m . The frame is called global if $U = M$. Similarly, a local n -frame is a collection of n pointwise linearly independent sections.

The following is an exercise.

Proposition 8.1.13 ([7, Prop. 10.15]). Suppose $\pi : E \rightarrow M$ is a smooth **VB** of rank k .

- (a) If $\{\sigma_i\}_{i=1}^r$ is a smooth local r -frame of E over an open set $U \subset M$, then there exist local sections $\sigma_{r+1}, \dots, \sigma_k$ on some neighborhood $V \subseteq U$ such that $\{\sigma_i\}_{i=1}^k$ is a local frame for E over V .
- (b) A local frame of E over a closed set $A \subseteq M$ that is smooth in the sense that each of the sections in it can be extended to an open neighborhood of any point in A , then it can be extended to a smooth local frame on an open neighborhood of A .

Example 8.1.14. 1. The trivial product bundle $M \times \mathbb{R}^k \rightarrow M$ has a global frame given by the constant sections $\tilde{e}_i(p) = (p, e_i)$ where e_i are the standard basis vectors of \mathbb{R}^n .
2. Each local frame is *associated* with a local trivialization. Given a local trivialization $\chi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^k$, we can easily define a local frame by $\sigma_i = \chi^{-1} \circ \tilde{e}_i$. The converse will be proven in the next Proposition.

Proposition 8.1.15 ([7, Prop. 10.19]). Every smooth local frame of a smooth **VB** is associated with a smooth local trivialization.

Proof. We need to construct a trivialization given a frame. Given a smooth local frame (σ_i) define a map $\Psi : U \times \mathbb{R}^k \rightarrow \pi^{-1}(U)$ by

$$\Psi(m, (v^1, \dots, v^k)) = \sum_{i=1}^k v^i \sigma_i(m). \quad (8.1.5)$$

The fact that $(\sigma_i(m))$ forms a basis of E_m at each $m \in U$ means that Ψ is bijective, and clearly $\sigma_i = \Psi \circ \tilde{\epsilon}_i$. Thus it remains to show that Ψ is a diffeomorphism, which will turn it into a local trivialization.

The reason Ψ is a diffeomorphism is that we already know that E is a smooth **VB**, so, by shrinking U if necessary, we can assume that it has a local trivializing diffeomorphism $\chi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^k$. The maps $\chi \circ \sigma_i : U \rightarrow U \times \mathbb{R}^k$ are smooth with smooth components $\text{id}_U \times \sigma_i^j(k)$, $j = 1, \dots, k$, and therefore the map

$$\chi \circ \Psi(m, (v^1, \dots, v^k)) = (m, (v^i \sigma_i^1(m), \dots, v^i \sigma_i^k(m))), \quad (8.1.6)$$

is clearly smooth. The matrix $\sigma_i^j(m)$ is invertible for each m , and since inversion is a smooth map on $\text{GL}_n(\mathbb{R})$, its inverse τ_i^j is also smooth function of m . Therefore

$$(\chi \circ \Psi)^{-1}(m, (w^1, \dots, w^k)) = (m, (w^i \tau_i^1(m), \dots, w^i \tau_i^k(m))) \quad (8.1.7)$$

is also smooth, proving that Ψ is a diffeomorphism. ■

Remark 8.1.16. Proposition 8.1.15 is the reason why local trivializations of general S -bundles are called frames (see Definitions 7.3.1 and 7.4.1).

Proposition 8.1.17 (Sections of tensor products). *Let E and E' be two **FBs** over the same base manifold M . Then there is a canonical isomorphism*

$$\Gamma^\infty(E \otimes E') \cong \Gamma^\infty(E) \otimes_{C^\infty(M)} \Gamma^\infty(E'). \quad (8.1.8)$$

Proof. For simplicity we will write Γ instead of Γ^∞ . Consider the map $\alpha : \Gamma(E) \times \Gamma(E') \rightarrow \Gamma(E \otimes E')$ defined by $\alpha(\sigma, \sigma')(m) = \sigma(m) \otimes \sigma'(m)$. Since this map is clearly $C^\infty(M)$ -bilinear, it can be interpreted as a map on $\Gamma(E) \otimes \Gamma(E')$. We need to show it is an isomorphism.

First if E and E' are trivial, then $\Gamma(E)$ and $\Gamma(E')$ are both free $C^\infty(M)$ modules because every section is a linear combination of a global frame of sections $\{e_i\}$ (or $\{e'_i\}$ for E') with coefficients that are smooth functions on M . Then a global frame on $E \otimes E'$ is $e_i \otimes e'_j$ and α is clearly an isomorphism.

Now, if E and E' are still trivial but $E = E_1 \oplus E_2$, then we have

$$\begin{aligned} (\Gamma(E_1) \otimes \Gamma(E')) \oplus (\Gamma(E_2) \otimes \Gamma(E')) &\cong (\Gamma(E_1) \oplus \Gamma(E_2)) \otimes \Gamma(E') \cong \\ &\cong \Gamma(E \otimes E') \cong \Gamma(E_1 \otimes E') \oplus \Gamma(E_2 \otimes E'), \end{aligned} \quad (8.1.9)$$

which implies that the statement holds for $E_1 \otimes E'$ and $E_2 \otimes E'$ separately. All isomorphisms here remain natural because we use nothing but the natural projections coming with the definitions of direct sums and tensor products.

Finally, we will independently prove in Theorem 8.5.6 that every vector bundle is a direct summand of a trivial bundle, which completes the proof. Alternatively, one can trivialize E and E' over an open covering of M , decompose sections using a subordinate POU and check that α is an isomorphism explicitly. ■

Definition 8.1.18 (Vector subbundle). A (immersed, weakly embedded, or embedded) subbundle of a smooth vector bundle $E \xrightarrow{\pi} M$ is a vector bundle $E' \xrightarrow{\pi'} M$ with an injective VB morphism $\tilde{h} : E' \rightarrow E$ which turns E' into an immersed, weakly embedded, or embedded, respectively, submanifold of E . If $M = M'$ then we usually assume that \tilde{h} covers the identity in the base, and we call such subbundles *vertical*.

Remark 8.1.19. Vertical subbundles are always embedded (since injective images of local frames in E' can be extended to a full local frame on E by Proposition 8.1.13).

Definition 8.1.20 (Distribution). A distribution $D \subset E$ on a vector bundle $E \xrightarrow{\pi} M$ is a choice of a linear subspace $D_m \subset E_m$ for every $m \in M$. (There are no smoothness restrictions, and the dimension of D_m can change arbitrarily with m .)

- Theorem 8.1.21.**
1. A VB is trivial iff it has a global frame.
 2. Subbundles are embedded submanifolds.
 3. A distribution D is a subbundle of E iff it has smooth local frames (i.e. collections of pointwise linearly independent local sections of E that span D in every fiber).
 4. For any VB morphism $f : E \rightarrow E'$ of constant rank on fibers (as a linear map on every fiber), the distributions $\ker f = \sqcup_m \ker f_m \subset E$ and $\text{Im } f = \sqcup_m \text{Im } f_m \subset E'$ are subbundles (here f_m is the restriction to the fiber E_m).
 5. Given a subbundle $E' \subset E$, the quotient space E/E' , defined as the disjoint union of fiberwise quotients of vector spaces, is a smooth vector bundle.

Proof. 1 and 2 are a simple exercise. 4 and 5 can be shown by explicitly constructing local frames, which is easy because locally all bundles are trivial, see [7, Thm. 10.34].

3. By Proposition 8.1.13 we can complete a smooth local frame $\{\sigma_i\}_{i=1}^r$ for D to a smooth local frame $\{\sigma_i\}_{i=1}^k$ for E over some $U \subset M$. This local frame is associated with a smooth local trivialization χ defined by

$$\chi(s^1 \sigma_1(m) + \cdots + s^k \sigma_k(m)) = (m, (s^1, \dots, s^k)). \quad (8.1.10)$$

This map takes D locally to the subset $\{(m, (s^1, \dots, s^m, 0, \dots, 0)) \subset U \times \mathbb{R}^k$, which is an embedded submanifold. By composing this map with the projection onto the first r components of \mathbb{R}^k we get a smooth local trivialization of D , making it a smooth **VB**. ■

Lemma 8.1.22 ([7, Lem. 10.29]). *Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M$ be smooth VBs. A map $\mathcal{F} : \Gamma^\infty(E) \rightarrow \Gamma^\infty(E')$ is linear over $C^\infty(M)$ iff there is a smooth bundle morphism $F : E \rightarrow E'$ over M such that $\mathcal{F}(\sigma) = F \circ \sigma$ for all $\sigma \in \Gamma^\infty(E)$.*

Proof. The map on sections induced by a bundle morphism is clearly $C^\infty(M)$ -linear. Conversely, given \mathcal{F} , first we show its locality: if a section vanishes locally on an open set $U \subset M$, then its image under \mathcal{F} does as well. This follows by constructing, for each $m \in U$, a local bump function χ_m supported in U that equals 1 at $m \in U$ and using $\mathcal{F}(\chi_m \sigma) = \chi_m \mathcal{F}(\sigma) = 0$, which, upon evaluation at m , gives $\mathcal{F}(\sigma)(m) = 0$.

Next we show that \mathcal{F} is pointwise (*hyperlocal*): if $\sigma(m) = 0$, then $\mathcal{F}(\sigma)(m) = 0$. This follows by constructing a local frame at m , extending it to a collection of global sections $s_i \in \Gamma(E)$, and replacing σ by $\sum_i c^i s_i$ with $c^i \in C^\infty(M)$ on a neighborhood of m . Then, by the locality property we just proved, $\mathcal{F}(\sigma)(m) = \sum_i \mathcal{F}(c^i s_i)(m) = \sum_i c^i(m) \mathcal{F}(s_i)(m) = 0$.

Now we can construct a bundle morphism F corresponding to \mathcal{F} . For any $m \in M$ and $e \in E_p$ let $F(e) = \mathcal{F}(s)(m) \in E'_m$, where $s \in \Gamma(E)$ is a global section such that $s(m) = e$. Hyperlocality implies that the resulting element of E'_m is independent of the choice of s . This map clearly satisfies $\pi' \circ F = \pi$, and is fiberwise linear because of linearity of \mathcal{F} . It also satisfies $F \circ \sigma(m) = \mathcal{F}(\sigma)(m)$ for all $\sigma \in \Gamma(E)$ by definition.

It remains to show the smoothness of F , which can be done locally. Given $m \in M$, let $s_i \in \Gamma(E)$ be a collection of sections that form a local frame on a neighborhood of m . We may also assume that a similar collection $s'_i \in \Gamma(E')$ exists on the same neighborhood. There exist local smooth functions $A_i^j \in C^\infty(U)$ such that $\mathcal{F}(s_i)|_U = \sum_j A_i^j s'_j$.

For any $m \in U$ and $e \in E_m$ we can write $e = e^i s_i(m)$ for some numbers $e^i \in \mathbb{R}$. Then

$$F(e^i s_i(m)) = \mathcal{F}(e^i s_i)(m) = e^i \mathcal{F}(s_i)(m) = e^i A_i^j(m) s'_j(m). \quad (8.1.11)$$

If χ and χ' denote the local trivializations of E and E' associated with the frames $\{s_i\}$ and $\{s'_i\}$, respectively, then the local representative of F with respect to these trivializations has the form

$$\chi' \circ F \circ \chi'(m, (e^1, \dots, e^k)) = (m, (A_i^1(m)e^i, \dots, A_i^{k'}(m)e^i)) \quad (8.1.12)$$

(summations over i implied), which is smooth. ■

Definition 8.1.23 (Quotient bundle). Let $L \subset E$ be a vertical subbundle of a **VB** $E \xrightarrow{\pi} M$. Let

$$E/L = \bigsqcup_{m \in M} E_m/L_m \quad (8.1.13)$$

and let $\pi' : E/L \rightarrow M$ be the natural projection to the index set. By construction the fibers of π' are vector spaces. There exists a family of local trivializations $\{(U_\alpha, \chi_\alpha)\}$ of E and a subspace $K \subset F$ of the typical fiber F of E such that the restrictions of χ_α to L take values in $U_\alpha \times K$. For any such χ_α define a mapping

$$\chi'_\alpha : (\pi')^{-1}(U_\alpha) \rightarrow U_\alpha \times F/K, \quad \chi'_\alpha([x]) = (\pi(x), [\chi_{\alpha,m}(x)]). \quad (8.1.14)$$

It is an exercise to check the differentiability of the corresponding transition maps. The resulting vector bundle $\pi' : E/L \rightarrow M$ is called the quotient bundle of E by L .

Remark 8.1.24. There is a way to avoid having to construct the quotient bundle manually out of fibers. As we will discuss in Section 18.1, every vector bundle admits a bundle metric, i.e. a smoothly varying choice of a dot product on each fiber. Then we can define the orthogonal complement L^\perp in the obvious way, and this will automatically be a smooth **VB** such that $L \oplus L^\perp \cong E$, and up to bundle isomorphism it will be independent of the choice of metric. Then we can define E/L up to isomorphism as the bundle L^\perp .

The following is an exercise.

Proposition 8.1.25. An open covering $\{U_\alpha\}$ of M and a cocycle $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \mathrm{GL}_k(\mathbb{R})$ define a unique **VB** up to isomorphism, which can be constructed as the quotient $E := (\bigsqcup_\alpha U_\alpha \times \mathbb{R}^k) / \sim$, where $(m_1, v_1) \sim (m_2, v_2)$ iff $m_1 = m_2 = m$ and $v_1 = \tau_{\beta\alpha}(m) \cdot v_2$ for the corresponding α, β .

Example 8.1.26 (Tautological bundles of Grassmannians). Let V be a topological vector space and define its Grassmannian $G_k(V)$ as the topological space whose points are the k -dimensional vector subspaces of V . It can be constructed as the quotient of the subset of the direct product $V^{\times k}$ consisting of k -tuples of linearly independent elements of V (i.e. k -frames in V) by the action of GL_k on it, as we will see in Section 13.8.

In the case $k = 1$ we get the projective space of V . When $V = \mathbb{K}^n$ for some field \mathbb{K} , we have the homeomorphisms $\mathbb{K}P^{n-1} \cong G_1(\mathbb{K}^n)$.

All Grassmannians of real or complex vector spaces are compact smooth manifolds, because it is easy to construct a finite smooth atlas consisting of precompact sets. This smooth structure will also follow from the quotient manifold theorems of Section 13.7.

The *tautological bundle* over $G_k(V)$ is defined as the rank k vector bundle $\gamma_k(V) \xrightarrow{\pi} G_k(V)$ in which the fiber above a point $L \in G_k(V)$ is the vector space L itself. In other words, the

total space consists of all pairs (L, v) with $L < V$ and $v \in L$. It inherits the subspace topology from $G_k(V) \times V$. The projection map is $\pi : (L, v) \mapsto L$. For example, the tautological vector bundle $\gamma_1(\mathbb{R}^2) \rightarrow \mathbb{RP}^1 \cong \mathbb{S}^1$ is smoothly isomorphic to the Möbius band bundle. We will revisit these bundles in detail in Section 13.8.

Exercise 8.1. Check that Grassmannians are smooth manifolds (by giving an atlas) and their tautological bundles are smooth vector bundles (by showing local triviality). See also the solution of Problem 2.22 in [9]. What is the dimension of $G_k(\mathbb{K}^n)$? Show that the tautological bundles are smooth VBs by constructing local frames.

Exercise 8.2. The projective complex line $\mathbb{C}P^1$ is diffeomorphic to the sphere \mathbb{S}^2 . Its tautological line bundle is a complex line (or real plane) bundle over the sphere. Show that it is exactly $\mathcal{O}(-1)$ (Example 8.1.5), and thus its circle bundle is exactly the Hopf bundle.

8.2 Tangent bundle

Definition 8.2.1 (Tangent bundle). Let M be a manifold with a smooth atlas $\{(U_\alpha, \varphi_\alpha)\}$. Its tangent bundle TM is defined as the disjoint union of tangent spaces as fibers $TM = \bigsqcup_{m \in M} T_m M$, with the structure of a smooth VB specified by the VB atlas consisting of the push-forwards $\{(TM|_{U_\alpha}, (\varphi_\alpha)_*)\}$. Indeed, since $\varphi_\alpha : U_\alpha \rightarrow \hat{U}_\alpha \subset \mathbb{R}^m$ are diffeomorphisms, we have diffeomorphisms $(\varphi_\alpha)_* : TM|_{U_\alpha} \rightarrow \hat{U}_\alpha \times \mathbb{R}^m$, which are indeed local trivializations for the bundle. The linear structure induced on fibers agrees with the one we have constructed previously on every tangent space.

Explicitly, a local trivialization represents a tangent vector (m, v) as $((x^1, \dots, x^m), v^i \partial_i)$, where $v^i \partial_i = \sum_{i=1}^m v^i \frac{\partial}{\partial x^i}$ is understood as an element of \mathbb{R}^m in the symbolic basis $\{\frac{\partial}{\partial x^i}\}$ of \mathbb{R}^m .

A local chart φ_α gives rise to a local frame (the pullback of the standard basis/frame on \mathbb{R}^m) conventionally denoted by partials,

$$\partial_i := \varphi_{\alpha*}^{-1} \left(\frac{\partial}{\partial x^i} \right). \quad (8.2.1)$$

Example 8.2.2 (Tangent bundle of \mathbb{S}^2). If we embed \mathbb{S}^2 into \mathbb{R}^3 in the standard way, then its tangent bundle can be written as

$$T\mathbb{S}^2 = \{(\mathbf{x}, \mathbf{v}) \in \mathbb{S}^2 \times \mathbb{R}^3 \mid \mathbf{x} \perp \mathbf{v}\}. \quad (8.2.2)$$

The easiest way to understand its bundle structure is by using the structure of a complex manifold on $\mathbb{C}P^1$ introduced in Example 7.7.8. Since the charts (U_0, z) and (U_∞, w) are related by $w(z) = \frac{1}{z}$, the transition function on the tangent bundle $T\mathbb{C}P^1$, which is a complex

line bundle, is

$$\tau_{0\infty}(z) = \frac{dz}{dw} = -\frac{1}{w^2} = -z^2. \quad (8.2.3)$$

The sign here can be removed by a rotation of a chart, therefore this bundle is isomorphic to $O(2)$ as defined in Example 8.1.5. In particular, they are also isomorphic as real plane bundles.

Since the structure group \mathbb{C}^\times deformation retracts onto U_1 (this necessarily breaks the analytic structure), we should be able to find an atlas with orthogonal transition functions. Representation (8.2.2) suggests exactly such an atlas. We express the stereographic coordinates z, w in terms of the spherical angles (θ, φ) on the unit sphere:

$$z = \tan \frac{\theta}{2} e^{i\varphi}, \quad w = \cot \frac{\theta}{2} e^{-i\varphi}. \quad (8.2.4)$$

Now the Jacobi matrix of the transition map $z(w)$ in these coordinates reads

$$(\varphi_{0\infty})_* = \frac{\partial z}{\partial w} = \frac{\partial z}{\partial(\theta, \varphi)} \cdot \left(\frac{\partial w}{\partial(\theta, \varphi)} \right)^{-1} = -\tan^2 \frac{\theta}{2} \begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ \sin 2\varphi & \cos 2\varphi \end{pmatrix}. \quad (8.2.5)$$

Obviously this is just the real form of $-z^2$. Let us write out the corresponding bundle charts. Decompose a tangent vector v at (θ, φ) in the coordinate basis, $v = v^\theta \partial_\theta + v^\varphi \partial_\varphi$, and compute the bundle charts with values in $T_z \mathbb{C} \cong T_w \mathbb{C} \cong \mathbb{C}$:

$$z_* \cdot \begin{pmatrix} v^\theta \\ v^\varphi \end{pmatrix} = z \cdot \left(\frac{v^\theta}{\sin \theta} + iv^\varphi \right) = \frac{e^{i\varphi}}{2 \cos^2 \frac{\theta}{2}} (v^\theta + i \sin \theta v^\varphi), \quad (8.2.6)$$

$$w_* \cdot \begin{pmatrix} v^\theta \\ v^\varphi \end{pmatrix} = -w \cdot \left(\frac{v^\theta}{\sin \theta} + iv^\varphi \right) = -\frac{e^{-i\varphi}}{2 \sin^2 \frac{\theta}{2}} (v^\theta + i \sin \theta v^\varphi). \quad (8.2.7)$$

Note that, despite the spherical coordinates being undefined at the poles, these two formulas can be smoothly extended to cover their respective poles. For example, at the north pole any tangent vector is the limit of vectors $v^\theta \partial_\theta$ as $\theta \rightarrow 0$ at fixed φ . Such a vector is then locally represented by $\frac{1}{2} v^\theta e^{i\varphi}$, and one can check that this is smooth extension of the bundle chart z_* . It is now clear that to reduce this atlas to an SO_2 -structure we only need to rescale these chart maps in each fiber by a factor of $\cos^2 \frac{\theta}{2}$ or $-\sin^2 \frac{\theta}{2}$, respectively:

$$\chi_0(\mathbf{x}, \mathbf{v}) := \left(z, \frac{e^{i\varphi}}{2} (v^\theta + iv^\varphi \sin \theta) \right), \quad \chi_\infty(\mathbf{x}, \mathbf{v}) = \left(w, \frac{e^{-i\varphi}}{2} (v^\theta + iv^\varphi \sin \theta) \right). \quad (8.2.8)$$

These charts still extend smoothly over their respective poles, and the transition function becomes $e^{2i\varphi}$. This atlas endows $T\mathbb{S}^2$ with an SO_2 -structure, including an orientation.

Definition 8.2.3 (Vector fields). Vector fields on M are sections of TM . The vector space of all smooth vector fields is denoted $\mathfrak{X}(M) := \Gamma^\infty(TM)$. In a local frame $\{e_i\}$ on $U \subset M$, every vector field can be expanded as $X|_U = \sum_i X^i e_i$ with components $X^i \in C^\infty(U)$.

Definition 8.2.4 (Parallelism). A smooth manifold M^m is called *parallelizable* if TM is a trivial vector bundle, i.e. if there exist m point-wise linearly independent vector fields on M . A specific choice of a global frame is called an *absolute parallelism*.

Remark 8.2.5. Each absolute parallelism determines a canonical flat connection, where the parallel transports of any vector leave its components with respect to the distinguished global frame constant. However, this is not the only connection naturally arising from an absolute parallelism. It is flat, but not torsion-free, so one can also define a connection with the opposite torsion, or a symmetric one with no torsion. Finally, an absolute parallelism also naturally determines a Riemannian metric $g = \sum_i \alpha_i \otimes \alpha_i$ (where $\alpha_i \in \Omega^1(M)$ is the 1-form dual to e_i), and thus there is a Levi-Civita connection associated with this metric, which doesn't generally coincide with any of the other natural connections mentioned above. In general, none of these connections besides the first one are flat.

Proposition 8.2.6 ([16, Prop. 2.3.17]). *The spheres \mathbb{S}^1 , \mathbb{S}^3 , and \mathbb{S}^7 are parallelizable.*

Proof. In the case of \mathbb{S}^1 we identify \mathbb{R}^2 with \mathbb{C} , so the Euclidean dot product can be written as $x \cdot y = \operatorname{Re}(\bar{x}y)$ and vector fields on \mathbb{S}^1 are represented by maps $X : \mathbb{S}^1 \rightarrow \mathbb{C}$ satisfying $\operatorname{Re}(\bar{z}X(z)) = 0$. This condition holds for example for $X(z) = iz$. Since this function is nowhere vanishing, the corresponding vector field forms a global frame for $T\mathbb{S}^1$.

In the case of \mathbb{S}^3 we identify \mathbb{R}^4 with \mathbb{H} and the Euclidean norm is again $x \cdot y = \operatorname{Re}(\bar{x}y)$, where \bar{x} now represents quaternionic conjugation. Vector fields on \mathbb{S}^3 are represented by maps $X : \mathbb{S}^3 \rightarrow \mathbb{H}$ such that $\operatorname{Re}(\bar{\mathbf{q}}X(\mathbf{q})) = 0$. For $l = 1, 2, 3$, define vector fields $X_l : \mathbb{H} \rightarrow \mathbb{H}$ by

$$X_1(\mathbf{q}) = \mathbf{q}\mathbf{i}, \quad X_2(\mathbf{q}) = \mathbf{q}\mathbf{j}, \quad X_3(\mathbf{q}) = \mathbf{q}\mathbf{k}. \quad (8.2.9)$$

Then $\operatorname{Re}(\bar{\mathbf{q}}X_l(\mathbf{q})) = 0$ and $\operatorname{Re}(\overline{X_l(\mathbf{q})}X_j(\mathbf{q})) = \delta_{lj}$. Hence the X_l restrict to vector fields on \mathbb{S}^3 and they are pointwise linearly independent. In the case of \mathbb{S}^7 , the proof is analogous, with quaternions replaced by octonions. ■

Remark 8.2.7. The proof of the above Proposition carries over to the unit spheres of \mathbb{C}^k , \mathbb{H}^k and \mathbb{O}^k , where \mathbb{O} denotes the octonions. Thus, for $r = 2, 4, 8$ and $k = 1, 2, \dots$, there exist $r - 1$ pointwise linearly independent vector fields on \mathbb{S}^{rk-1} . In case $k = 1$, these vector fields constitute a global frame, whereas in the other cases just a global $(r - 1)$ -frame. Establishing whether these $(r - 1)$ -frames extend to full frames on odd-dimensional spheres is non-trivial. On the other hand, on an even-dimensional sphere, every vector field has a zero (Hairy Ball Theorem).

The question of the parallelizability of the spheres \mathbb{S}^n was open until the early 20th century when the Hurwitz-Radon-Eckmann theorem established the following. Let $n \in \mathbb{N}$, then n can be uniquely written in the form $n = k \cdot 2^\nu$ with k odd, $\nu = 4b + c$, $0 \leq c \leq 3$. Define $\rho(n) = 8b + 2^c$ (independent of k). Then the sphere \mathbb{S}^{n-1} has $\rho(n) - 1$ linearly independent vector fields and there is no larger global frame.

2^ν	1	2	4	8	16	32	64	128
$\rho(n)$	1	2	4	8	9	10	12	16

(8.2.10)

In particular, this implies that the only parallelizable spheres are the ones that are Lie groups (in fact, all Lie groups are parallelizable), $\mathbb{S}^0 = O(1)$, $\mathbb{S}^1 = U(1)$, $\mathbb{S}^3 = SU(2)$, plus the exceptional \mathbb{S}^7 .²⁹

The proof of this theorem includes an explicit construction of the linearly independent vector fields using Clifford algebras and then a difficult proof of the non-existence of a larger set of independent vector fields using K-theory. This proof became the first big success of K-theory.

Definition 8.2.8 (Tangent functor). Define the functor $T : \text{Man}^\infty \rightarrow \text{VectBun}^\infty$ by $T : M \mapsto TM$ and

$$T : C^\infty(M, N) \rightarrow C^\infty(TM, TN), \quad f \mapsto Tf := f_*. \quad (8.2.11)$$

In local trivializations $\{(\varphi_\alpha)_*\}$ and $\{(\psi_\beta)_*\}$ of TM and TN , respectively, we have the local representatives $\widehat{f}_* = (\psi_\beta)_* \circ f_* \circ (\varphi_\alpha)^{-1}_* = (\psi_\beta \circ f \circ \varphi_\alpha^{-1})_* = (\widehat{f})_*$, which are smooth as derivatives of a smooth maps acting $U_\alpha \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^n$. This is also a **VB** morphism because it is linear on fibers.

Proposition 8.2.9. *The following properties make sure that T is a covariant functor.*

1. Identity property of a functor: $T \text{id}_M = \text{id}_{TM}$
2. T is a covariant functor, $T(f \circ g) = Tf \circ Tg$;
3. For a diffeomorphism f , $T(f^{-1}) = (Tf)^{-1}$.

Proof. Follows from Theorem 3.6.6. ■

Definition 8.2.10 (Lie derivatives). The Lie derivative of a function $f \in C^\infty(M)$ along a vector field $X \in \mathfrak{X}(M)$ is defined as the function

$$\mathcal{L}_X f \equiv X \cdot f := df(X) \in C^\infty(M), \quad (8.2.12)$$

where $df = \text{pr}_2 \circ f_*$ with $f_* : TM \rightarrow T\mathbb{R}$ and pr_2 being the projection onto the second factor (the fiber) in $T\mathbb{R} \cong \mathbb{R} \times \mathbb{R}$.

- Clearly, $(df)(X)(p) = (d_p f)(X_p) \in \mathbb{R}$ and $\mathcal{L}_X f \in C^\infty(M)$ because both f_* and X are smooth. In local coordinates $x^i(p) = \varphi_\alpha^i(p)$, if $X = \sum X^i \partial_i$, then

$$\mathcal{L}_X f(p) = X^i \partial_i f(p) = X^i \mathcal{L}_{\partial_i} f(p) = X^i \frac{\partial}{\partial x^i} f(\varphi_\alpha^{-1}(x)), \quad (8.2.13)$$

i.e. just a directional derivative along the vector field.

- A *diffeomorphism* $F : M \rightarrow N$ gives rise to a diffeomorphism $F_* : TM \rightarrow TN$, and it can act on vector fields as $F_* : \mathfrak{X}(M) \rightarrow \mathfrak{X}(N)$. Lie derivatives are natural w.r.t. push-forwards:

$$\mathcal{L}_{F_* X} f = df \circ F_* \circ X = d(f \circ F) \circ X = \mathcal{L}_X (f \circ F) = \mathcal{L}_X (F^* f). \quad (8.2.14)$$

- \mathcal{L}_X is a derivation on the algebra $C^\infty(M)$ of functions and $\mathcal{L}_X \cdot 1 = 0$.

Definition 8.2.11 (Cotangent bundle). Using the Hom-functor defined in Proposition 8.1.10, we define the cotangent bundle of a manifold M by

$$T^* M = \text{Hom}(TM, M \times \mathbb{R}), \quad (8.2.15)$$

where $\text{pr}_1 : M \times \mathbb{R} \rightarrow M$ is the trivial product bundle. The fiber $T_m^* M$ of the cotangent bundle can be naturally identified with the dual of the corresponding fiber $T_m M$. If $\tau_{\alpha\beta}$ are the transition functions on TM , then the transition functions on $T^* M$ are the inverse transposes, $(\tau_{\alpha\beta}^T)^{-1}$. Elements of the cotangent bundle are called *cotangent vectors*, or simply *covectors* on M .

We will return to a more detailed study of the cotangent bundle in the next section.

Another important class of bundles, called jet bundles, are a generalization of the tangent and cotangent bundles. Recall that one way of defining tangent vectors was as directional derivatives acting on smooth functions. Similarly covectors in $T_p^* M$ are linear functionals on vectors, hence they can be identified with differentials $d_p f$ of smooth functions. The differential $d_p f$ can be identified with the linear term of the Taylor series for f at p in any local coordinate system. Thinking of functions $C^\infty(M)$ as sections of the trivial bundle $M \times \mathbb{R}$, we can alternatively describe covectors at $p \in M$ as equivalence classes of (local) sections of this bundle where two sections are considered equivalent if the order 1 terms in their Taylor expansions at $p \in M$ in any local coordinate chart coincide. Extending this to higher derivatives naturally leads to the following construction.

Definition 8.2.12 (Germs, jet bundles, prolongations). 1. Given two smooth manifolds M, N and a point $m \in M$ define the equivalence relation \sim_p on $C^\infty(M, N)$ by saying

that $f \sim_p g$ iff there exists an open neighborhood $U_p \subset M$ of p such that $f|_{U_p} = g|_{U_p}$. The equivalence class of a given map f under this relation is called the germ of f at p , denoted $[f]_p$. The space of germs is denoted

$$C_p^\infty(M, N) = C^\infty(M, N)/\sim_p . \quad (8.2.16)$$

If $f_1 \in C^\infty(M, N)$ and $f_2 \in C^\infty(N, K)$, then we can compose germs as follows:

$$[f_2]_{f_1(p)} \circ [f_1]_p = [f_2 \circ f_1]_p. \quad (8.2.17)$$

2. For any $k \in \mathbb{N}$, another equivalence relation on $C^\infty(M, N)$ is defined as follows: $f \sim_{p;k} g$ iff in any, and therefore all, local coordinate chart, the local representatives \hat{f} and \hat{g} have Taylor series that coincide up to and including all terms of order k . The equivalence class of a given map f under this relation is called the k -jet of f at p , denoted $j_p^k f$. The set of all k -jets at p is denoted

$$J_p^k(M, N) = C^\infty(M, N)/\sim_{p;k} . \quad (8.2.18)$$

The composition of germs induces compositions of k -jets:

$$(j_{f_1(p)}^k f_2) \circ (j_p^k f_1) = j_p^k(f_2 \circ f_1). \quad (8.2.19)$$

A jet is called invertible if there is a jet $j_{f(p)}^k g \in J_{f(p)}^k(M, M)$ such that $j_p^k(g \circ f) = j_p^k \text{id}_M$ and $j_{f(p)}^k(f \circ g) = j_{f(p)}^k \text{id}_M$.

On the sets of all k -jets $J^k(M, N) = \bigsqcup_{p \in M} J_p^k(M, N)$ we can define natural projections

$$\pi_l^k : J^k(M, N) \rightarrow J^l(M, N), \quad j_p^l f \mapsto p, \quad 0 \leq l < k, \quad (8.2.20)$$

$$\pi_0^k : J^k(M, N) \rightarrow M \times N, \quad j_p^k f \mapsto (p, f(p)). \quad (8.2.21)$$

The projections π_0^k induce a unique smooth structure on $J^k(M, N)$ that turns both of the projections into smooth fiber bundles. The map π_0^k is called the k -th jet bundle of maps from M to N . We set

$$J_p^k(M, N)_q = (\pi_0^k)^{-1}((p, q)). \quad (8.2.22)$$

If $\dim M = m$ and $\dim N = n$, then the typical fiber of the jet bundle is the manifold $J_0^k(\mathbb{R}^m, \mathbb{R}^n)_0$, which parametrizes the values of Taylor coefficients (above the 0-th order) of maps in a local chart.

3. If $\pi : E \rightarrow M$ is a smooth **FB**, then the set of k -jets of smooth local sections of E is denoted $J^k(E)$ and is called the *jet bundle of order k* of E . It turns out to be an

embedded submanifold of $J^k(M, E)$ and a smooth **FB** over M . Clearly J^k is a functor acting on locally invertible bundle morphisms, called the *jet prolongation functor*. Given a local section $\sigma \in \Gamma_U^\infty(E)$ for $U \subset M$, its k -th prolongation is the induced section $j^k\sigma \in \Gamma_U^\infty(J^k(E))$ defined by

$$(j^k\sigma)(p) = j_p^k\sigma. \quad (8.2.23)$$

Note that j^k is a differential operator. Local sections of $J^k(E)$ that have this form for some local section $\sigma \in \Gamma^\infty(E)$ are called *holonomic*.

4. For every operator $D : \Gamma^\infty(E \rightarrow M) \rightarrow \Gamma^\infty(E' \rightarrow M)$ that is a k -th order differential operator in local coordinates, there exists a bundle morphism $\tilde{D} : J^k(E) \rightarrow E'$ such that $D = \tilde{D} \circ j^k$.
5. An order k *partial differential equation* on E is a closed embedded submanifold S of the jet bundle $J_M^k(E)$. A solution is a local section $\sigma \in \Gamma_U(J^k(E))$ that takes values in S .
6. The structure group of an order k jet bundle of maps on an m -dimensional manifold is called the *k -th differential group*

$$G_m^k := J_{0,0}^k(\mathbb{R}^m, \mathbb{R}^m)^\times. \quad (8.2.24)$$

Example 8.2.13. 1. Consider the trivial bundle $\text{pr}_1 : M \times \mathbb{R} \rightarrow M$. Then there is a bundle isomorphism between $J^1(\text{pr}_1) \rightarrow M$ and the bundle $T^*M \times \mathbb{R} \rightarrow M$ (whose fibers are $T_p^*M \times \mathbb{R}$). Indeed, every section $\sigma : M \rightarrow M \times \mathbb{R}$ can be identified, by projecting onto the second component, with a smooth function $f : M \rightarrow \mathbb{R}$. Then $j_p^1\sigma$ can be identified with the element $(df, f(p)) \in T^*M \times \mathbb{R}$.

2. Similarly, considering the trivial bundle $\text{pr}_2 : M \times \mathbb{R} \rightarrow \mathbb{R}$, its first jet bundle $J^1(\text{pr}_2)$ can be identified with $TM \times \mathbb{R} \rightarrow \mathbb{R}$. Every section $\sigma : \mathbb{R} \rightarrow M \times \mathbb{R}$ can be identified with a smooth path $\gamma : \mathbb{R} \rightarrow M$, and a first jet of such paths is a tangent vector by definition, so $j_t^1\sigma$ can be identified with $\dot{\gamma}(t) \in T_{\gamma(t)}M$.
3. In other words, $TM \cong J_0^1(\mathbb{R}, M)$ and $T^*M \cong J^1(M, \mathbb{R})_0$. The k -th order analog of the tangent bundle is $J_0^k(\mathbb{R}, M)$, whose typical fiber is the set $J_0^k(\mathbb{R}, \mathbb{R}^m)$ of the first k derivatives of paths in \mathbb{R}^m .

8.3 Tensor bundles

Recall from Example 1.5.11 that the space $\text{Hom}^k(E_1, \dots, E_k; F)$ of k -linear maps $E_1 \times \dots \times E_k \rightarrow F$, where E_i and F are vector spaces, can be naturally identified with the space $\text{Hom}(E_1 \otimes \dots \otimes E_k, F)$ of linear maps $E_1 \otimes \dots \otimes E_k \rightarrow F$. This hom-functor, as we know from Example 1.2.3, is contravariant with respect to the E_i 's and covariant with respect to F .

In particular, the dual of a real vector space E is defined as $E^* = \text{Hom}(E, \mathbb{R})$, and is a

contravariant functor of E . If the elements of E are called vectors, then the elements of E^* are called *covectors*. Being a functor, dualization acts on linear maps $f \in \text{Hom}(E, F) \mapsto f^* \in \text{Hom}(F^*, E^*)$ as $f^*(\alpha)(v) = \alpha(f(v))$, $v \in E, \alpha \in F^*$. We will sometimes use the notation $\langle \alpha, v \rangle := \alpha(v)$ for the natural pairing of covectors and vectors. Then $\langle f^*\alpha, v \rangle = \langle \alpha, fv \rangle$.

Moreover, we know from Example 1.5.15 that there is a natural isomorphism of vector spaces

$$\text{Hom}(E_1 \otimes \cdots \otimes E_r \otimes \cdots \otimes E_k, F) \cong \text{Hom}(E_1 \otimes \cdots \otimes E_r, \text{Hom}(E_{r+1} \otimes \cdots \otimes E_k, F)). \quad (8.3.1)$$

The particular case of $F = \mathbb{R}$ leads to

$$(E_1 \otimes \cdots \otimes E_r \otimes \cdots \otimes E_k)^* \cong \text{Hom}(E_1 \otimes \cdots \otimes E_r, E_{r+1}^* \otimes \cdots \otimes E_k^*). \quad (8.3.2)$$

With this in mind, we first define the tensor vector spaces.

Definition 8.3.1 (Tensors). For a real vector space E , the space of tensors of type (r, s) (r is called the contravariant rank and s is called the covariant rank) on E is the space of multilinear functionals of s vectors and r covectors,

$$\mathbb{T}_s^r(E) = \text{Hom}^{r+s}(\underbrace{E^*, \dots, E^*}_r, \underbrace{E, \dots, E}_s; \mathbb{R}) \cong \text{Hom}\left((E^*)^{\otimes r} \otimes E^{\otimes s}, \mathbb{R}\right) \cong E^{\otimes r} \otimes (E^*)^{\otimes s}. \quad (8.3.3)$$

We also have the extra isomorphisms

$$\mathbb{T}_s^r(E) \cong \text{Hom}(E^{\otimes s}, E^{\otimes r}) \cong \text{Hom}\left((E^*)^{\otimes r}, (E^*)^{\otimes s}\right). \quad (8.3.4)$$

In other words, a tensor of type (r, s) is a multilinear map that takes in s vectors and outputs r “vectors” (more precisely, linear combinations of tensor products of r vectors). Alternatively, it takes in r covectors and outputs s covectors.

Definition 8.3.2 (Tensor functors). We define the (partial) functors $\mathbb{T}_s^r : \text{FinVect}_{\mathbb{R}} \rightarrow \text{FinVect}_{\mathbb{R}}$ by their action on linear maps $f : E \rightarrow F$. Namely, for a tensor $t \in \mathbb{T}_s^r(E)$, we define $(\mathbb{T}_s^r f)(t) \in \mathbb{T}_s^r(F)$ by

$$(\mathbb{T}_s^r f)(t)(\alpha^1, \dots, \alpha^r; v_1, \dots, v_s) = t(f^*\alpha^1, \dots, f^*\alpha^r; f^{-1}v_1, \dots, f^{-1}v_s), \quad \alpha_i \in E^*, v_j \in E. \quad (8.3.5)$$

We denote $f_s^r := \mathbb{T}_s^r f$. It is very important to note that this is well-defined only when either t is totally contravariant ($s = 0$) or when f is invertible. Therefore \mathbb{T}_0^r is a true functor on the full category of vector spaces, whereas \mathbb{T}_s^r for $s > 0$ are functors only on the category of isomorphisms.

Remark 8.3.3. The terminology of co- and contra-varianc tensors is unfortunate: \mathbb{T}_0^1 is a covariant functor but produces the space of “contravariant” tensors, etc.

- Example 8.3.4.**
1. $\mathbb{T}_0^1(E) = \text{Hom}(E^*, \mathbb{R}) = E^{**} \cong E$ (for finite-dimensional E , which is the only case we consider) – vectors are contravariant tensors of rank 1; $f_0^1 = f$;
 2. $\mathbb{T}_1^0(E) = \text{Hom}(E, \mathbb{R}) = E^*$ – covectors are covariant tensors of rank 1;
 3. $f_0^1 = f$ and $f_1^0 = (f^{-1})^*$. There is no way to compare components of f and f^* as matrices in any basis since they don't act on the same vector spaces, but $\sum_j \left((f^{-1})^* \right)_i^j \cdot (f)_j^k = \delta_i^k$.
 4. $\mathbb{T}_1^1(E) \cong \text{End}(E)$ – the space of linear operators on E , i.e. square matrices are tensors of type $(1, 1)$;
 5. $\mathbb{T}_2^0(E) \cong \text{Hom}(E, E^*)$ – covariant tensors of rank 2 (examples are moment of inertia tensors, Riemannian metrics, stress tensors). They convert vectors to covectors.
 6. $\mathbb{T}_0^2(E) \cong \text{Hom}(E^*, E)$ are tensors that convert covectors to vectors.

Definition 8.3.5 (Components of tensors). Let $\{e_i\}_{i=1}^n$ be a basis of E and $\{\alpha^i\}_{i=1}^n$ the dual basis in E^* . The components of a tensor $t \in \mathbb{T}_s^r E$ in these bases are the set of n^{r+s} numbers defined as

$$t_{i_1 \dots i_s}^{j_1 \dots j_r} = t(\alpha^{j_1}, \dots, \alpha^{j_r}; e_{i_1}, \dots, e_{i_s}). \quad (8.3.6)$$

If a linear map $f : E \rightarrow E'$ in some bases $\{e_i\}, \{e'_j\}$ is represented by a matrix $A_i^j = \alpha'^j(f(e_i))$, then

$$((\mathbb{T}_s^r f)(t))_{\bullet \bullet \bullet}^{\bullet \bullet \bullet} = A_{j_1}^{\bullet} \cdots A_{j_r}^{\bullet} \left(A^{-1} \right)_{\bullet}^{i_1} \cdots \left(A^{-1} \right)_{\bullet}^{i_s} t_{i_1 \dots i_s}^{j_1 \dots j_r}, \quad (8.3.7)$$

where we use the Einstein summation convention and \bullet denotes unnamed free indices.

Now we can use the procedure in Definition 8.1.10 to extend the tensor functors to bundles.

Definition 8.3.6 (Tensor bundles). Given a vector bundle $E \xrightarrow{\pi} M$, we define its tensor powers $\mathbb{T}_s^r(E)$ by applying the \mathbb{T}_s^r functor fiberwise. In particular, a **VB** atlas on $\mathbb{T}_s^r E$ can be generated as $(\chi_\alpha)_s^r$ from a **VB** atlas of local trivializations χ_α on E . We have the **VB** isomorphisms

$$\mathbb{T}_s^r(E) \cong E^{\otimes r} \otimes (E^*)^{\otimes s}. \quad (8.3.8)$$

In addition, we define tensor bundles of a manifold M as

$$T_s^r M := \mathbb{T}_s^r(TM) \cong (TM)^{\otimes r} \otimes (T^*M)^{\otimes s}, \quad T_{\bullet}^{\bullet} M := \bigoplus_{r,s=0}^{\infty} T_s^r M. \quad (8.3.9)$$

In particular, $T^*M = T_1^0 M$ is the *cotangent bundle* of M . We also define for $f \in C^\infty(M, N)$

$$T_s^r f := (f_*)_s^r : T_s^r M \rightarrow T_s^r N. \quad (8.3.10)$$

Any local chart φ_α on M generates a local trivialization $T_s^r \phi_\alpha$ of the bundle $T_s^r M$. The spaces of *tensor fields* on M are

$$\Gamma_s^r(M) := \Gamma^\infty(T_s^r M). \quad (8.3.11)$$

In particular, $\mathfrak{X}(M) = \Gamma_0^1(M)$ is the space of vector fields and $\mathfrak{X}^*(M) = \Gamma_1^0(M)$ is the space of *covector fields*, also known as *differential one-forms*. The spaces of tensor fields are $C^\infty(M)$ -modules because we can multiply them by scalar functions fiberwise as

$$(f \cdot t)(p) = f(p) \cdot t(p). \quad (8.3.12)$$

Obviously, this is nothing but a special case of the tensor product of tensor fields

$$t \in \Gamma_{s_1}^{r_1}, t' \in \Gamma_{s_2}^{r_2} \implies t \otimes t' \in \Gamma_{s_1+s_2}^{r_1+r_2} \quad (8.3.13)$$

defined also fiberwise. The components of a tensor product of tensors are simply products of the components:

$$(t \otimes t')_{j_1 \dots j_{s_1+s_2}}^{i_1 \dots i_{r_1+r_2}} = t_{j_1 \dots j_{s_1}}^{i_1 \dots i_{r_1}} \cdot (t')_{j_{s_1+1} \dots j_{s_1+s_2}}^{i_{r_1+1} \dots i_{r_1+r_2}}. \quad (8.3.14)$$

With this product, the total tensor algebra is

$$\Gamma_\bullet^r(M) = \bigoplus_{r,s=0}^{\infty} \Gamma_s^r(M). \quad (8.3.15)$$

Corollary 8.3.7. *From Example 1.5.15 and Proposition 8.1.17, we have the isomorphism of $C^\infty(M)$ -modules*

$$\Gamma_s^r(M) \cong (\mathfrak{X}(M))^{\otimes r} \otimes_{C^\infty(M)} (\mathfrak{X}^*(M))^{\otimes s}. \quad (8.3.16)$$

Remark 8.3.8. This corollary reveals, perhaps, the most fundamental reason for why differential geometry involves a radically new level of algebra compared to analysis on \mathbb{R}^n , where all we need is linear algebra over the field of real numbers. As we see, in geometry we have to work with linear algebra not over a field, but over the ring of functions $C^\infty(M)$ (not to mention that this ring is “infinite-dimensional”). The structure of the appearing algebraic objects like spaces of tensor fields ends up depending not only on the local structure of the manifold (which is identical to that of \mathbb{R}^n , where we can get away with just algebra over fields), but also on the global, often topological, properties of the manifold. Eventually this

will force us to develop a whole new algebraic formalism that is able to process ring modules instead of vector spaces, called *homological algebra*. The tools coming with it will let us extract knowledge about the global topology of manifolds and bundles just by studying the structure of the associated algebraic objects.

Theorem 8.3.9 (Tensor characterization). *There is a natural isomorphism of $C^\infty(M)$ -modules*

$$\Gamma_s^r(M) \cong \text{Hom}_{C^\infty(M)}(\mathfrak{X}^*(M)^{\otimes r} \otimes \mathfrak{X}(M)^{\otimes s}, C^\infty(M)). \quad (8.3.17)$$

The right-hand side here is the space of $C^\infty(M)$ -valued functionals of a number of vector and covector fields that are multilinear over $C^\infty(M)$. This means that tensors are not just ultralocal functionals (depend only on local values of the input fields), but their values at a point in fact depend only on the values of the input fields at that very point (not on any of their derivatives etc.).

Proof. **WLOG**, let $r = 0$, $t \in \Gamma_s^0(M)$ and $f \in C^\infty(M)$. Clearly t is a C^∞ -linear functional of fields:

$$\begin{aligned} (t(fX_1, \dots, X_s))(p) &= t(p)(f(p)X_1(p), \dots, X_s(p)) = \\ &= f(p)t(p)(X_1(p), \dots, X_s(p)) = ((f \cdot t)(X_1, \dots, X_s))(p). \end{aligned} \quad (8.3.18)$$

Conversely, given such a functional \mathcal{A} , we need to show that it defines a covariant tensor. First, $C^\infty(M)$ -linearity implies locality: if ψ is a bump function around $p \in M$, then

$$\mathcal{A}(\psi \cdot X_1, X_2, \dots, X_s)(p) = \psi(p)\mathcal{A}(\dots)(p), \quad (8.3.19)$$

therefore the value of $\mathcal{A}(\dots)$ at p depends only on $X_i(p)$.

Now let us pick a local frame $\{\partial_i\}$ and arbitrarily extend it to a set of global vector fields $\{E_i\}$ such that $E_i(p) = \partial_i(p)$. Since we have already established locality, it suffices to consider vector fields of the form $X_i = X_i^j E_j$ with $X_i^j \in C^\infty(M)$. Then

$$\begin{aligned} \mathcal{A}(X_1, \dots, X_s)(p) &= X_1^j(p)\mathcal{A}(E_j, X_2, \dots, X_s)(p) = \dots \\ &\dots = X_1^{j_1} \cdots X_s^{j_s}(p) \underbrace{\mathcal{A}(E_{j_1}, \dots, E_{j_s})(p)}_{\in \mathbb{R}}, \end{aligned} \quad (8.3.20)$$

which indeed depends only on $X_i^j(p)$. It also implies that if $X_i(p) = Y_i(p)$, then $\mathcal{A}(\{X\})(p) = \mathcal{A}(\{Y\})(p)$.

Let $t_p(v_1, \dots, v_s) = \mathcal{A}(X_1, \dots, X_s)(p)$ with $X_i(p) = v_i$ for each $p \in M$. Furthermore, the components of t locally are $t_{i_1 \dots i_s} = \mathcal{A}(\partial_{i_1}, \dots, \partial_{i_s})$, which are smooth by assumption. This

means that t constructed this way is a smooth section of the tensor bundle, i.e. a tensor field. ■

Remark 8.3.10. This isomorphism is extremely important in proving that various objects which naively look like differential operators, are in fact tensor fields. The Riemann curvature tensor, which is naively a second order differential operator (commutator of two covariant derivatives), is a canonical example of such a tensor, but we will encounter more examples much sooner than that.

Example 8.3.11 ($\mathbf{d}f$ is a one-form). For any scalar function $f \in C^\infty(M)$, its differential $\mathbf{d}f$ defined in Def. (8.2.10) can be viewed as a functional from smooth vector fields to smooth functions because $\mathbf{d}f(X) = \mathbf{d}f \circ X \in C^\infty(M)$ for any $X \in \mathfrak{X}(M)$ and it is $C^\infty(M)$ -linear as a functional because $\mathbf{d}f(g \cdot X) = g \cdot \mathbf{d}f(X)$ for all $g \in C^\infty(M)$. Therefore $\mathbf{d}f \in \mathfrak{X}^*(M)$ by the above Theorem.

Definition 8.3.12 (Push-forward, Pullback). Let $f \in C^\infty(M, N)$ be a *diffeomorphism* and $t \in \Gamma_s^r(M)$. Then we extend the push-forward to all tensors by

$$f_* t = (Tf)_s^r \circ t \circ f^{-1} = (f_*)_s^r \circ t \circ f^{-1}. \quad (8.3.21)$$

Similarly, for a tensor field $t \in \Gamma_s^r(N)$ we define the pullback

$$f^* t = \left(f^{-1}\right)_* t = \left(f_*^{-1}\right)_s^r \circ t \circ f. \quad (8.3.22)$$

Note that both of these definitions use f^{-1} , so they cannot be extended to non-diffeomorphisms in general. However, in the case of pullbacks of totally covariant tensors ($r = 0$) we can bypass this obstruction. Recall that $T_s^0 f = f_*^{-1} \otimes \cdots \otimes f_*^{-1}$, which means that

$$\text{for covariant } t \in \Gamma_s^0(N) : \quad f^* t = (f_*)^{\otimes s} \circ t \circ f, \quad (8.3.23)$$

or in terms of its action on vector fields,

$$(f^* t)(X_1, \dots, X_s) = t(f_* X_1, \dots, f_* X_s). \quad (8.3.24)$$

Here, $f_* X_i$ are the pointwise push-forwards of tangent vectors. Push-forwards of contravariant tensors are always well-defined pointwise, but then only form a global smooth tensor on the target manifold if f is a diffeomorphism.

Remark 8.3.13. From properties of T -functors, we have

$$(f \circ g)_* = f_* \circ g_*, \quad (f \circ g)^* = g^* \circ f^*. \quad (8.3.25)$$

8.4 Whitney approximation and embedding

In many cases it is helpful to think of abstract manifolds as submanifolds of some Euclidean space of a sufficiently high dimension. The following theorem shows that this is indeed possible for any manifold.

Theorem 8.4.1 (Whitney Embedding Theorems). *Any m -dimensional smooth manifold M can be*

1. *smoothly embedded into a Euclidean space of a sufficiently high dimension;*
2. *smoothly immersed in \mathbb{R}^{2m} (sharp version – into \mathbb{R}^{2m-1});*
3. *smoothly properly embedded in \mathbb{R}^{2m+1} (sharp version – into \mathbb{R}^{2m} , but not properly);*

Proof. The idea of the proof is to use charts to map pieces of M into \mathbb{R}^m and then use the extra dimensions to “bend” the image of each subsequent chart into a new dimension to avoid self-overlaps. Finally, one needs to reduce the final dimension as much as possible (the hard part).

Let us only provide a simple proof of (1) in the compact case. If M is compact, it has a finite covering by coordinate balls $\{(B_i, \varphi_i)\}_{i=1}^n$. Constructing bump functions χ_i that are 1 on B_i , we can define

$$F(p) = (\chi_1\varphi_1(p), \dots, \chi_n\varphi_n(p)) \in \mathbb{R}^{nm}. \quad (8.4.1)$$

It is easy to check that this is a smooth immersion and, since M is compact, an embedding. See the general proof in [7, Thm. 6.15]. ■

Recall that a manifold is said to have a compatible analytic structure if its differentiable structure contains an atlas in which all transition functions are real analytic (i.e. have Taylor series converging in a non-zero radius at every point). Since the standard smooth structure of \mathbb{R}^n is obviously analytic, we get the following.

Corollary 8.4.2. *Every smooth manifold has a compatible structure of a real analytic manifold.*

Remark 8.4.3. The Whitney Embedding Theorem, in fact, holds for all C^1 -manifolds (in the compact case the proof we gave above suffices). Thus every C^1 -manifold can be made analytic. Another sophisticated theorem of Morrey and Grauert shows that the analytic structure on a (paracompact) smooth manifold is actually unique.

In most applications we will need only the first (the weakest) version of the embedding theorem.

Theorem 8.4.4 (Smooth Approximation Theorem [7, Cor. 6.22]). *Any continuous map $f \in C(M, \mathbb{R}^k)$ has a smooth δ -close smooth approximation $\tilde{f} \in C^\infty(M, \mathbb{R}^k)$ for any $\delta \in C(M, (0, +\infty))$ (meaning $|f(p) - \tilde{f}(p)| < \delta(p)$ for all $p \in M$).*

Proof. The idea is to cover the manifold by a countable, locally finite open covering $\{U_i\}$ whose parts are so small that the values of f stay within $\delta(p_i)$ from each other on U_i , where p_i is an arbitrarily fixed point in each U_i . Then taking a POU χ_i subordinate to the cover, we can define

$$\tilde{f} = \sum_i \chi_i \cdot f(p_i). \quad (8.4.2)$$

It is obvious that $|f - \tilde{f}| < \delta$ everywhere. ■

Definition 8.4.5 (Normal bundle). If an m -dimensional manifold M is embedded into some \mathbb{R}^n , we define its normal bundle fiberwise as $N_p M = (T_p M)^\perp$, where \perp is the orthogonal complement of a subspace taken with respect to the standard dot product on \mathbb{R}^n . Then NM is an n -dimensional embedded submanifold of $T\mathbb{R}^n$.

Definition 8.4.6 (Tubular neighborhood). A tubular neighborhood of a manifold $M \lhd \mathbb{R}^n$ is its open neighborhood $M \subset U$ such that U is diffeomorphic to the normal bundle NM .

Theorem 8.4.7 (Tubular neighborhood theorem [7, Thm. 6.24]). *Every embedded $M \lhd \mathbb{R}^n$ has a tubular neighborhood.*

Proof. Thinking of NM as an embedded submanifold of $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$, we can consider the function $g \in C^\infty(NM, \mathbb{R}^n)$ defined as

$$g(x, v) = x + v. \quad (8.4.3)$$

It maps the fiber $N_x M$ to the subspace through x that is orthogonal to $T_x M$ (the image of g is literally the union of all hyperplanes of dimension $n - \dim M$ orthogonal to M). The zero section NM is diffeomorphic to M , and g can be shown to be a local diffeomorphism on a neighborhood of the zero section. This then implies that there is a neighborhood of the zero section within NM on which g is a diffeomorphism between a neighborhood of M and a neighborhood of a zero section in NM . But any neighborhood of a zero section of a vector bundle is diffeomorphic to the entire bundle. See the details of the proof in [7, Thm. 6.24]. ■

Remark 8.4.8. 1. The Tubular Neighborhood theorem generalizes to embedded submanifolds of arbitrary smooth manifolds, not just \mathbb{R}^n . Instead of straight lines one then considers geodesics with respect to an arbitrary Riemannian metric. Alternatively, one can use the Whitney Embedding Theorem.

2. When M is non-compact, the “thickness” of the tubular neighborhood may have to vary with the point $m \in M$ and may not be bounded from below by a positive constant.

Theorem 8.4.9. Any tubular neighborhood U of $M \lhd \mathbb{R}^n$ can be retracted onto M by a smooth submersion, i.e. $\exists \pi : U \rightarrow M$ such that $\pi|_M = \text{id}_M$ and π is a smooth submersion.

Proof. Just define $\pi = \pi_{NM} \circ g^{-1}$ with π_{NM} the bundle projection of the normal bundle and g from the last proof. ■

Theorem 8.4.10 (Whitney Approximation Theorem). Any continuous map $f \in C(N, M)$ between two smooth manifolds is homotopic to a smooth map.

Proof. Embed $M \lhd \mathbb{R}^n$, let U be a tubular neighborhood, $\pi : U \rightarrow M$ the corresponding retraction and $\delta(x) = \sup\{\epsilon \leq 1 \mid B_\epsilon(x) \subset U\} \in C(M, (0, +\infty))$.

Then $\tilde{\delta} = \delta \circ f \in C(N, (0, +\infty))$ and there exists a smooth approximation $\tilde{f} \in C^\infty(N, \mathbb{R}^n)$ that is $\tilde{\delta}$ -close to f . But now the image of \tilde{f} is not in M , but in U !

We can fix this by defining the homotopy

$$H(n, t) = \pi \underbrace{\left((1-t)f(n) + t\tilde{f}(n) \right)}_{\in U}. \quad (8.4.4)$$

This is a homotopy between f and $\pi \circ \tilde{f}$. ■

8.5 Transversality

Definition 8.5.1 (Transversality). Let $S, S' \lhd M$ be two embedded smooth manifolds. They are said to *intersect transversally* at $p \in M$ if $T_p M = T_p S + T_p S'$. More generally, a map $f : N \rightarrow M$ is said to be *transverse* to $S \lhd M$ if $T_{f(n)} S + f_{*n}(T_n N) = T_{f(n)} M$. Finally, two maps $f : N \rightarrow M$ and $g : N' \rightarrow M$ are said to be transverse if $f_{*n}(T_n N) + g_{*n'}(T_{n'} N') = T_p M$ for every $p \in \text{im}(f) \cap \text{im}(g)$ and n, n' such that $f(n) = g(n') = p$. The notation is $S \pitchfork S'$, $f \pitchfork S$, and $f \pitchfork g$.

Theorem 8.5.2 ([7, Thm. 6.30]). 1. If $F \in C^\infty(N, M)$ is a smooth map transverse to an embedded submanifold $S \lhd M$, then $F^{-1}(S)$ is an embedded submanifold of N of the same codimension as S in M .
2. If $S, S' \lhd M$ intersect only transversally, then $S \cap S' \lhd M$ is an embedded submanifold of codimension $\text{codim}(S \cap S') = \text{codim } S + \text{codim } S'$.

Proof. The second statement follows from the first by taking F to be the inclusion $S' \hookrightarrow M$ and noting that a composition of smooth embeddings $S \cap S' \hookrightarrow S \hookrightarrow M$ is again a smooth embedding.

Let $m = \dim M$ and $k = \text{codim } S$. Given $x \in F^{-1}(S)$, we can find a neighborhood U of $F(x)$ in M and a local defining function $\varphi : U \rightarrow \mathbb{R}^k$ for S , so that $S \cap U = \varphi^{-1}(0)$. The theorem will be proven if we show that 0 is a regular value of $\varphi \circ F$ because $F^{-1}(S) \cap F^{-1}(U)$ is the zero set of $\varphi \circ F|_{F^{-1}(U)}$.

To show that 0 is a regular value we need to show that the differential of $\varphi \circ F$ is surjective at every point $p \in (\varphi \circ F)^{-1}(0)$. But φ is constant along S and F is transverse to it, which means that the rank of $\varphi \circ F$ matches the codimension of S , which is k , making p a regular point. Thus $F^{-1}(S)$ is an embedded submanifold of codimension k . ■

For example, in \mathbb{R}^3 , this theorem shows that a smooth curve and a smooth surface intersecting transversely have only isolated points in their intersection, while two smooth surfaces intersect transversely in a smooth curve. Two smooth curves in \mathbb{R}^3 intersect transversally iff their intersection is empty.

Corollary 8.5.3. *If $F : N \rightarrow M$ is a submersion and $S \lhd M$ an embedded submanifold, then $F^{-1}(S)$ is an embedded submanifold of the same codimension.*

Proof. Every submersion is transverse to every embedded submanifold. ■

The key fact about transversely intersecting submanifolds and transverse maps is that they are “generic” as we will now show. The key to finding transverse maps is the following application of Sard’s theorem, which gives a simple sufficient condition for a family of smooth maps to contain at least one map that is transverse to a given submanifold.

Theorem 8.5.4 (Parametric Transversality Theorem [7, Thm. 6.35]). *Let $F : N \times D \rightarrow M$ be a smooth family of maps $F_s : N \rightarrow M$, $s \in D$ (D is connected), and $S \lhd M$ an embedded submanifold. If $F \pitchfork S$, then $F_s \pitchfork S$ for almost every $s \in D$.*

Proof. The hypothesis implies that $W = F^{-1}(S)$ is an embedded submanifold of $N \times D$ by Theorem 8.5.2. Let $\pi : N \times D \rightarrow D$ be the projection onto the second factor. What we will show is that if $s \in D$ is a regular value of $\pi|_W$, then $F_s \pitchfork S$. Since almost every s is a regular value by Sard’s Theorem 3.4.4, this finishes the proof.

Suppose $s \in D$ is a regular value of $\pi|_W$. Pick a $p \in F_s^{-1}(S)$ and let $q = F_s(p) \in S$. We need to show that $T_q M = T_q S + F_{s*}(T_p N)$. By the hypothesis on F , we already know that $T_q M = T_q S + F_{s*}(T_{(p,s)}(N \times D))$. Since s is a regular value and $(p, s) \in W$, we also know

$T_s D = \pi_*(T_{(p,s)} W)$. Transversality also implies $F_*(T_{(p,s)} W) = T_q S$. Putting all of this together,

$$\begin{aligned} T_q M &= T_q S + F_*(T_{(p,s)}(N \times D)) = T_q S + F_*(T_p N \oplus \pi_*(T_{(p,s)} W)) = \\ &= T_q S + F_{s*} \circ \pi_*(T_p N) + F_{p*}(T_{(p,s)} W), \end{aligned} \quad (8.5.1)$$

where $F_p(s) = F(p, s)$. But $F_{p*} \circ \pi_*(T_{(p,s)} W)$ is merely a restriction of F_* , so its image is already inside $T_q S$, thus in fact $T_q M = T_q S + F_{s*}(T_p N)$, which proves $F_s \pitchfork S$. ■

Theorem 8.5.5 (Transversality homotopy theorem [7, Thm. 6.36]). *Any map $f \in C^\infty(N, M)$ is homotopic to a $g \in C^\infty(N, M)$ that is transverse to a given embedded submanifold $S \lhd M$.*

Proof. The crux of the proof is constructing a smooth map $F : N \times D \rightarrow M$ that is transverse to S , where D is the ball \mathbb{D}^k for some k and $F_0 = f$. It then follows from Theorem 8.5.4 that there is some $s \in D$ such that $F_s : N \rightarrow M$ is transverse to S , and that F_s is homotopic to f (since D is path-connected).

By the Whitney Embedding Theorem 8.4.1, we can assume that M is properly embedded into some \mathbb{R}^k . Let U be a tubular neighborhood of M in \mathbb{R}^k and let $r : U \rightarrow M$ be a smooth retraction that is also a smooth submersion. If we define $\delta : M \rightarrow \mathbb{R}_+$ by $\delta(x) = \sup\{\epsilon \leq 1 : B_\epsilon(x) \subset U\}$, as in the Whitney Approximation Theorem 8.4.10, then Theorem 8.4.4 shows that there exists a smooth function $e : N \rightarrow \mathbb{R}_+$ that satisfies $0 < e(p) < \delta(f(p))$ everywhere. Let D be the unit ball in \mathbb{R}^k and define $F : N \times D \rightarrow M$ by

$$F(p, s) = r(f(p) + e(p) \cdot s). \quad (8.5.2)$$

Clearly F is smooth and $F_0 = f$ because r is a retraction.

For each $p \in N$, the restriction of F to $\{p\} \times D$ is the composition of the local diffeomorphisms $s \mapsto f(p) + e(p)s$ followed by the smooth submersion r , making F a smooth submersion and hence transverse to S . ■

Theorem 8.5.6 ([19, Thm. II.14.2]). *For any VB $E \xrightarrow{\pi} M$ there exists a VB $E' \xrightarrow{\pi'} M$ such that $E \oplus E'$ is trivial.*

Proof. Embed a neighborhood $U \subset E$ of the zero section of E into \mathbb{R}^n by Whitney embedding. Let τ_U and ν_U be the tangent and the normal bundles of $U \subset \mathbb{R}^n$, respectively. Then $\tau_U \oplus \nu_U$ is trivial over U .

Vectors along the fibers of E become tangent vectors to U with nonzero projections onto $(T_x M)^\perp \subset T_x U$. This projection provides an isomorphism $E \cong N_U M \subset TU$ (by $N_U M$ we

mean the normal bundle of M as a manifold embedded into $U \subset \mathbb{R}^n$, not straight into \mathbb{R}^n). But $N_U M \oplus (\nu_U|_M) \cong \nu_M$ (where ν_M is the normal bundle of M taken in \mathbb{R}^n). Therefore $N_U M \oplus (\nu_U|_M) \oplus \tau_M \cong \nu_M \oplus \tau_M$, which is trivial! We can take $E' = \nu_U|_M \oplus \tau_M$. ■

Remark 8.5.7. The above theorem holds more broadly for continuous vector bundles over base spaces that are homotopy equivalent to a finite-dimensional CW-complex. All CW-complexes are paracompact, but paracompactness alone is not enough to guarantee this property. For example it fails for $M = \mathbb{R}P^\infty$.

Theorem 8.5.8 (Transverse section theorem [19, Thm. II.15.3]). *Let $E \xrightarrow{\pi} M$ be a smooth vector bundle and $U_0 \subset E$ an arbitrary open neighborhood of the zero section $E_0 \subset U_0$. Let M' be a smooth manifold and $f \in C^\infty(M', E)$. Then there exists a smooth section $\sigma : M \rightarrow E$ whose image lies in U_0 such that $f \pitchfork \sigma$.*

Proof. By Theorem 8.5.6, there is a vector bundle $P \rightarrow M$ such that $E \oplus P$ is trivial. Then we have the commutative diagram

$$\begin{array}{ccc} f^*(E \oplus P) & \xrightarrow{f'} & E \oplus P \\ \pi' \downarrow & \curvearrowright & \downarrow \Pi \\ M' & \xrightarrow{f} & E \end{array} \quad (8.5.3)$$

defining the map f' . A global trivialization of $E \oplus P$ provides a diffeomorphism $\phi : M \times \mathbb{R}^n \rightarrow E \oplus P$. Let $g : E \oplus P \rightarrow \mathbb{R}^n$ be the resulting projection onto the fiber. Let $z \in \mathbb{R}^k$ be a fixed regular value of $g \circ f' : f^*(E \oplus P) \rightarrow \mathbb{R}^n$. That is, $g_* \circ f'_*$ maps the tangent space at any point $v \in f^*(E \oplus P)$ with $g \circ f'(v) = z$ onto $T_z \mathbb{R}^n$. This implies that $\text{im } f'_*$ spans the complement of $T_{(f'(v), z)}(M \times \{z\})$ and therefore f' is transverse to the section $\sigma' : M \rightarrow E \oplus P$ given in terms of the trivialization by $\sigma'(p) = (p, z)$. Define the section $\sigma : M \rightarrow E$ by $\sigma(p) = \Pi \circ \sigma'(p)$, so that the following diagram commutes:

$$\begin{array}{ccccc} f^*(E \oplus P) & \xrightarrow{f'} & E \oplus P & \xleftarrow{\sigma'} & M \\ \pi' \downarrow & & \downarrow \Pi & \nearrow \sigma & \\ M' & \xrightarrow{f} & E & & \end{array} \quad (8.5.4)$$

We claim that $f \pitchfork \sigma$. To see this, let $x \in M'$ and $p \in M$ be such that $f(x) = \sigma(p)$. Then $\Pi \circ \sigma'(p) = \sigma(p) = f(x)$. By definition of the pullback bundle, the point $(x, \sigma'(p)) \in f^*(E \oplus P)$ has $f'((x, \sigma'(p))) = \sigma'(p)$. Since $f' \pitchfork \sigma'$, the images of the differentials

$f'_{*(x, \sigma'(p))}$ and σ'_{*p} span the tangent space $T_{f'((x, \sigma'(p)))=\sigma'(p)}(E \oplus P)$. Since Π is a submersion, this must also hold for the images of the differentials f_{*x} and σ'_{*p} to the tangent space $T_{f(x)=\sigma(p)}E$. This precisely means that $f \pitchfork \sigma$.

Clearly σ can be made arbitrarily close to the zero section because the neighborhood U_0 necessarily contains another neighborhood $E_0 \subset U \subset U_0$ such that $E \cong U$ as bundles, and then the proof still applies to U viewed as a bundle, with f restricted to $f : f^{-1}(U) \rightarrow U$ (here $f^{-1}(U)$ is a smooth manifold being an open subset of M'). ■

Corollary 8.5.9 (Zeros of generic sections of VBs). *Let $E \xrightarrow{\pi} M$ be a VB of rank k , and $\dim M = m$. Then:*

1. *If $k > m$, then there exists a section σ that is nowhere zero (doesn't intersect with the zero section). Moreover, by induction there exists a global $(k-m)$ -frame, which spans a trivial subbundle of rank $k-m$. We say that a "generic section" of this bundle has no zeros.*
2. *If $k = m$ (like $E = TM$), then the generic section turns to zero at most on a discrete subset of M ($\dim(\text{im}(\sigma) \cap E_0) = 0$). Any non-generic section can be deformed by an arbitrarily small homotopy into one.*
3. *If $k < m$, then the generic section turns to zero along a submanifold of M of dimension at most $m-k$ ($\dim(\text{im}(\sigma) \cap E_0) = m-k$).*

Proof. We prove claim 1, the rest are trivial via a similar argument. Take the zero section $\sigma_0 : M \rightarrow E$ and find a section σ transversal to it by Theorem 8.5.8. Then $\text{codim}(\text{im}(\sigma) \cap E_0) = 2k > m+k$, which means that this intersection is empty (two submanifolds whose dimensions add up to less than the dimension of the ambient manifold can only be transverse if they are disjoint; e.g. the intersection of two curves in \mathbb{R}^3 is generically empty). This theorem is due to J.P. Serre [18, Thm. 2]. ■

8.6 K-theory

We will denote a trivial vector bundle of rank k by ε^k . We only care about its isomorphism class, so we can safely assume the typical fiber is \mathbb{K}^k . In the following discussion the base manifold M is fixed.

Definition 8.6.1 (Stable isomorphism of VBs). *We say that E' and E'' are *stably isomorphic* ($E' \sim^s E''$) if there exists a trivial bundle ε^n such that $E' \oplus \varepsilon^n \cong E'' \oplus \varepsilon^n$. In particular, a bundle is called *stably trivial* iff its direct sum with some trivial bundle is trivial.*

Another equivalence relation which we will denote simply \sim is defined by the existence of two trivial bundles ε^k and ε^l such that $E' \oplus \varepsilon^k \cong E'' \oplus \varepsilon^l$.

Example 8.6.2. Stably trivial bundles don't have to be trivial. Consider the (non-trivial) tangent bundle of the sphere, $T\mathbb{S}^n$, and the normal line bundle $N\mathbb{S}^n$ of the embedding $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. Then $T\mathbb{S}^n$ is stably trivial because the direct sum $T\mathbb{S}^n \oplus N\mathbb{S}^n$ is trivial and $N\mathbb{S}^n$ is trivial.

More broadly, one can show that the question of the triviality of the normal bundle of an embedding $M \hookrightarrow \mathbb{R}^n$ for sufficiently high n does not depend on the embedding. Thus, if M admits an embedding into an \mathbb{R}^n with a trivial normal bundle, there is a well defined equivalence class called a *stably trivial normal bundle*.

Corollary 8.6.3 (Existence of \sim -inverses). *If for a given VB E there are two VBs E' and E'' such that $E \oplus E'$ and $E \oplus E''$ are both trivial, then*

$$\underbrace{(E \oplus E')}_{\text{trivial}} \oplus E'' \cong \underbrace{(E \oplus E'')}_{\text{trivial}} \oplus E'. \quad (8.6.1)$$

Therefore any two bundles that complete E to a trivial bundle are \sim -equivalent. Moreover, if E', E'' have the same rank, then they are also stably isomorphic.

This means that if $E \oplus E'$ is trivial, then \sim -equivalence class of E' , which exists by Theorem 8.5.6, can be seen as the “inverse” of the corresponding class of E in a certain group defined as follows.

Definition 8.6.4 (Reduced K-groups). The set of \sim -equivalence classes of real VBs over M forms a group under the Whitney sum \oplus . The inverse is defined as $[E] \oplus [E]^{-1} = [\text{trivial}]$. This group is denoted by $\widetilde{\text{KO}}(M)$ – the *reduced real K-group*. K is a functor, and this is the fundamental object of K-theory, and is itself an important topological invariant of manifolds. In the case of complex VBs the same definition produces the complex K-group $\widetilde{\text{K}}(M)$.

Example 8.6.5 (Bott periodicity). The famous Bott periodicity can be stated in terms of K-groups of spheres:

$n \bmod 8$	1	2	3	4	5	6	7	8
$\widetilde{\text{KO}}(\mathbb{S}^n)$	\mathbb{Z}_2	\mathbb{Z}_2	0	\mathbb{Z}	0	0	0	\mathbb{Z}

(8.6.2)

The first column is this table confirms what we've discovered in Example 7.5.3: up to addition of trivial bundles, there is only one non-trivial vector bundle over a circle – the Möbius band ($\text{GL}_n(\mathbb{R})$ has two connected components). However, this also encodes the information that the direct sum of two Möbius bands is a trivial plane bundle (any plane bundle over the circle is trivial). In general the proof of the Bott periodicity by Corollary 7.7.4 comes down to the computation of the homotopy groups of the orthogonal matrix groups, $\pi_{n-1}(\text{O}_k)$ for sufficiently large k (since $\text{GL}_k(\mathbb{R})$ is homotopy equivalent to O_k). This is a very deep result

originally obtained by Milnor via Morse theory. One finds that $\pi_{n+8}(O_\infty) = \pi_n(O_\infty)$. More generally, one can show that $\widetilde{KO}(M) \cong \widetilde{KO}(\Sigma^8 M)$ where Σ is the suspension functor. An even stronger source of this statement is the homotopy equivalence $\Omega^8 O_\infty \simeq O_\infty$, where Ω^8 is the eight-fold loop space functor.

For complex vector bundles Bott periodicity ends up being simpler, with a period of 2:

$n \bmod 2$	1	2
$\widetilde{K}(\mathbb{S}^n)$	0	\mathbb{Z}

(8.6.3)

which is a special case of the isomorphism $K(M) \cong K(\Sigma^2 M)$, and also follows from the equality $\pi_{n+2}(U_\infty) = \pi_n(U_\infty)$, or the homotopy equivalence $\Omega^2 U_\infty \simeq U_\infty$.

The stable isomorphism relation \sim_s doesn't immediately produce a group because it fixes the rank of the bundle, thus so far only the zero element $[\varepsilon^0]$ can have an inverse: $E \oplus E' \sim_s \varepsilon^0$ implies $E \oplus E' \oplus \varepsilon^n \cong \varepsilon^n$ for some n , which can only happen if E and E' have rank 0. However, even though inverses do not exist, Theorem 8.5.6 still gives us the cancellation property of addition:

Corollary 8.6.6 (Cancellation property). *If $E \oplus E_1 \sim_s E \oplus E_2$, then $E_1 \sim_s E_2$ (by adding to both sides a bundle E' such that $E \oplus E'$ is trivial).*

Recall that the group \mathbb{Z} of integers can be defined as equivalence classes of pairs of natural numbers $(m, n) \in \mathbb{N}^2$ under the equivalence relation $(m, n) \sim (m', n')$ iff $m + n' = m' + n$ with neutral element $[(0, 0)]$. The inverse of $[(m, n)]$ is $[(n, m)]$ and therefore we can identify the equivalence class $[(m, n)]$ with the difference $m - n$. Notably, every element of \mathbb{N} can be represented as $m = [(m, 0)]$, and every negative integer as $-m = [(0, m)]$. This construction is called the *Grothendieck group* and can be applied to any abelian monoid, such as a cancellative semigroup. Therefore we can use this to turn the stable isomorphism classes of **VBs** into an honest group by considering pairs, called “formal differences”.

Definition 8.6.7 (K -groups). The abelian groups $K(M)$ and $KO(M)$ consist of “formal differences” $E - E'$ of complex or real vector bundles over M , respectively, which are equivalence classes of pairs (E, E') under the equivalence relation $(E_1, E'_1) \sim (E_2, E'_2)$ iff $E_1 \oplus E'_2 \sim_s E_2 \oplus E'_1$. The transitivity of this relation follows from the cancellation property of Corollary 8.6.6. The addition of pairs is component-wise, the zero element is $[(E, E)]$, and the inverse of $[(E, E')]$ is $[(E', E)]$.

Let us focus on $K(M)$ for simplicity, although everything below works equally well for real vector bundles as well. Unlike with integers, every element of $K(M)$ can be represented as a difference $E - \varepsilon^n$ since if we start with $E - E'$ we can add to both E and E' a bundle which

trivializes E' . There is a natural homomorphism $K(M) \rightarrow \tilde{K}(M)$ sending $E - \varepsilon^n$ to the \sim -class of E . This is well-defined since if $E - \varepsilon^n = E' - \varepsilon^m$ in $K(M)$, then $E \oplus \overset{s}{\varepsilon^m} \sim E' \oplus \varepsilon^n$, hence $E \sim E'$. The map $K(M) \rightarrow \tilde{K}(M)$ is obviously surjective, and its kernel consists of elements $E - \varepsilon^n$ with $E \sim \varepsilon^0$, hence $E \sim \overset{s}{\varepsilon^m}$ for some m , so the kernel consists of elements of the form $\varepsilon^m - \varepsilon^n$. This kernel subgroup of $K(M)$ is isomorphic to \mathbb{Z} . In fact, restriction of **VBS** to a basepoint $x_0 \in M$ defines a homomorphism $K(M) \rightarrow K(x_0) \cong \mathbb{Z}$ which restricts to an isomorphism on the subgroup $\{\varepsilon^m - \varepsilon^n\}$. Thus we have a non-natural splitting dependent on x_0

$$K(M) \cong \tilde{K}(M) \oplus \mathbb{Z}, \quad (8.6.4)$$

which is why $\tilde{K}(M)$ is called the reduced group.

In addition to the additive group structure induced by the Whitney (direct) sum, $K(M)$ will also carry a ring structure induced by the Whitney (tensor) product. The product of elements of $K(M)$ is naturally defined as

$$(E_1 - E'_1)(E_2 - E'_2) = E_1 \otimes E_2 + E'_1 \otimes E'_2 - E_1 \otimes E_2 - E'_1 \otimes E_2. \quad (8.6.5)$$

It is routine to verify that this is well-defined and turns $K(M)$ into a commutative ring with identity ε^1 , the trivial line bundle (or the equivalence class $\varepsilon^1 - \varepsilon^0$ to be precise). We can simplify the notation by writing the element $\varepsilon^n \in K(M)$ as just n . This is consistent with arithmetic rules, e.g. nE is the sum of n copies of E .

If we choose $x_0 \in M$, then the map $K(M) \rightarrow K(x_0)$ obtained above is a ring homomorphism. Its kernel, which can be identified with $\tilde{K}(M)$ as we have seen, is an ideal, hence also a ring in its own right, although not necessarily with an identity.

Definition 8.6.8 (K-rings). With the multiplicative structure described above, \tilde{K} , \widetilde{KO} , K , and KO are contravariant functors into the category of commutative rings, with maps acting via pullbacks of bundles.

Definition 8.6.9 (External product). An external product $\mu : K(M) \otimes K(N) \rightarrow K(M \times N)$ can be defined by $\mu(a \otimes b) = \text{pr}_1^*(a) \text{pr}_2^*(b)$, where pr_1, pr_2 are the natural projections of $M \times N$. This is a ring homomorphism due to the commutativity of K-rings. This definition works similarly for the other three K-functors.

We can now easily show one direction of the isomorphism asserted by the complex Bott periodicity theorem. Recall that isomorphism classes of complex line bundles over $\mathbb{S}^2 = \mathbb{C}P^1$ are classified by a single integer. The tautological line bundle $\gamma_1(\mathbb{C}^2) \rightarrow \mathbb{C}P^1$ (whose unit circle bundle is the Hopf bundle) generates the rest of line bundles via tensor products. Let us denote this line bundle H (for Hopf). We now show that it satisfies a certain relation as an element of the K-ring of $\mathbb{C}P^1$. As before, we denote the trivial complex line bundle ε^1 by 1.

Proposition 8.6.10. *The line bundle H over $\mathbb{C}P^1$ satisfies $(H \otimes H) \oplus 1 \cong H \oplus H$.*

Proof. It suffices to provide a homotopy between the clutching functions. The clutching function of H is $\varphi(z) = z$ (defined in a neighborhood of $\mathbb{S}^1 \subset \mathbb{C}$), thus the clutching functions for $H \oplus H$ and $(H \otimes H) \oplus 1$ respectively are

$$\varphi_{H \oplus H}(z) = \begin{pmatrix} z & 0 \\ 0 & z \end{pmatrix}, \quad \varphi_{(H \otimes H) \oplus 1}(z) = \begin{pmatrix} z^2 & 0 \\ 0 & 1 \end{pmatrix}. \quad (8.6.6)$$

Now we rely on the fact that $\mathrm{GL}_n(\mathbb{C})$ is path-connected to say that there exists a continuous path $\gamma : [0, 1] \rightarrow \mathrm{GL}_2(\mathbb{C})$ such that $\gamma(0) = I$ and $\gamma(1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then a holomorphic homotopy between the two clutching functions written above is given by

$$g : [0, 1] \times \mathbb{S}^1 \rightarrow \mathrm{GL}_2(\mathbb{C}), \quad (t, z) \mapsto \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} \gamma(t) \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix} \gamma(t). \quad (8.6.7)$$

■

The above relation can be written as $H^2 + 1 = 2H$, and so $(H - 1)^2 = 0$.

Corollary 8.6.11. *There is a natural ring homomorphism $\mathbb{Z}[H]/(H - 1)^2 \rightarrow \mathrm{K}(\mathbb{C}P^1)$, and thus also a ring homomorphism*

$$\mathrm{K}(M) \otimes \mathbb{Z}[H]/(H - 1)^2 \rightarrow \mathrm{K}(M) \otimes \mathrm{K}(\mathbb{C}P^1) \xrightarrow{\mu} \mathrm{K}(M \times \mathbb{C}P^1). \quad (8.6.8)$$

The bulk of the proof of Bott periodicity for the K-functors then consists of showing that the maps

$$\mu : \mathrm{K}(M) \otimes \mathrm{K}(\mathbb{S}^2) \rightarrow \mathrm{K}(M \times \mathbb{S}^2), \quad \mu_0 : \mathrm{KO}(M) \otimes \mathrm{KO}(\mathbb{S}^8) \rightarrow \mathrm{KO}(M \times \mathbb{S}^8), \quad (8.6.9)$$

are isomorphisms. The following theorem is the main ingredient in complex Bott periodicity, which we state without proof. It can be proven via complex analytic methods, as in Hatcher, or via more general topological arguments, as we will do later.

Theorem 8.6.12 (Fundamental Product Theorem). *The homomorphism $\mathrm{K}(M) \otimes \mathbb{Z}[H]/(H - 1)^2 \rightarrow \mathrm{K}(M \times \mathbb{S}^2)$ is an isomorphism.*

One final homotopy invariant coming from the algebra of vector bundles derives from the obvious fact that a tensor product of two line bundles is still a line bundle, and every line bundle has a natural inverse – the dual bundle.

Definition 8.6.13 (Picard group). The set of isomorphism classes of line bundles over a manifold M forms a group under the tensor product \otimes . The identity element is the trivial line bundle $M \times \mathbb{K}$, and the inverse of a bundle is the dual bundle (because $E \otimes E^*$ has transition functions $\tau_{\alpha\beta} \otimes \tau_{\alpha\beta}^{-1} = \tau_{\alpha\beta} \tau_{\alpha\beta}^{-1} = 1$ since they are scalars for line bundles). This abelian group is called the Picard group of the manifold, $\text{Pic}(M)$.

Remark 8.6.14. The Picard group is especially important for complex manifolds, where it is the group of complex line bundles (with fiber \mathbb{C}). For instance, the Picard group of a Riemann surface is a discrete group whose elements are enumerated by their first Chern number. More generally, for nice spaces (say, smooth manifolds or CW-complexes) the Picard group is isomorphic to the second integral cohomology group $H^2(M, \mathbb{Z})$, and the isomorphism is provided by the first Chern class.

9 Vector Fields and Lie Derivatives

9.1 Vector fields and flows

Definition 9.1.1 (Integral curves). Let M be a smooth manifold and $X \in \mathfrak{X}(M)$ a vector field on it. A curve $\gamma : I \rightarrow M$ with $I = (a, b) \subset \mathbb{R}$ is an integral curve of X if the tangent vectors defined by γ on its domain coincide with the values of X :

$$\gamma'(t) = X(\gamma(t)), \quad t \in I. \quad (9.1.1)$$

In local coordinates, if γ is represented by $x^i(t)$, this is a system of (nonlinear) ODEs

$$\dot{x}^i(t) = X^i(x(t)). \quad (9.1.2)$$

Definition 9.1.2 (Flow boxes). Let $X \in \mathfrak{X}(M)$. A flow box of X at $p \in M$ is the triple (U_0, a, F) , where $U_0 \subset M$ is an open neighborhood of p , $a \in \mathbb{R}_{>0} \cup \{\infty\}$, and $F : U_0 \times I_a \rightarrow M$ with $I_a = (-a, a)$ is a smooth mapping, such that:

1. For all $p \in U_0$, the map $\gamma_p : I_a \rightarrow M$ defined by $\gamma_p(t) = F(p, t)$ is an integral curve for X at p ;
2. $F_t : U_0 \rightarrow M$ defined by $F_t(p) = F(p, t)$ is a diffeomorphism onto its image.

From this definition, we have the basic properties:

- $F_{t_1+t_2} = F_{t_1} \circ F_{t_2} = F_{t_2} \circ F_{t_1}, \quad F_0 = \text{id}, \quad (9.1.3)$
- On the common domain, $F_t^{-1} = F_{-t}$;
- Any vector field X has a flow box at every point $p \in \overset{\circ}{M}$. This follows from the results on the local solvability of systems of ODEs and smooth dependence of the solutions on the initial conditions. Since this dependence on the initial conditions has the identity Jacobi matrix at $t = 0$, F_t is a local diffeo and F_{-t} is the inverse.
- If F_t is the flow of X , then

$$F_{t*}X = X, \quad (9.1.4)$$

i.e. $F_{t*}(X(p)) = X(F_t(p))$. Moreover, X is the only vector field that satisfies this property for all (p, t) from an open neighborhood of $M \times \{0\}$ in $M \times \mathbb{R}$.

Definition 9.1.3 (Flow). The (maximal) flow of $X \in \mathfrak{X}(M)$ is a map $F : \mathcal{D} \rightarrow M$ defined on the (automatically open) set $\mathcal{D} \subset M \times \mathbb{R}$ that is maximal among all sets such that $(p, t) \in \mathcal{D}$ iff there is an integral curve γ_p of X such that $\gamma_p(0) = p$ and $F(p, t) = \gamma_p(t)$ for all $t \in \mathcal{D} \cap (\{p\} \times \mathbb{R})$. \mathcal{D} is called the *flow domain* of X . We also denote $\mathcal{D}_p = \mathcal{D} \cap (\{p\} \times \mathbb{R})$,

$\mathcal{D}_t = \mathcal{D} \cap (M \times \{t\})$, and $F_p : \mathcal{D}_p \rightarrow M$ is given by $F_p(t) = F(p, t) = \gamma_p(t)$.

Definition 9.1.4 (Complete vector field). $X \in \mathfrak{X}(M)$ is called complete if its flow domain is $\mathcal{D} = M \times \mathbb{R}$.

- Example 9.1.5.**
1. On $M = \mathbb{R}^2 \setminus \{0\}$, the vector field $X = \partial_x$ is incomplete: points on $\mathbb{R} \times \{0\}$ “fall off” the manifold in finite time (namely, $(-x, 0)$ runs into the non-existent point $(0, 0)$ at time $t = x$);
 2. On $M = \mathbb{R}$, the flow of the vector field $X = x^2 \partial_x$ is also incomplete because it evolves the point $x = 1$ according to $\gamma(t) = 1/(1-t)$ (solution of the ODE $\dot{x}(t) = x^2$), which runs away to infinity in finite time.
 3. Complete vector fields don’t form a linear space. Consider on \mathbb{R}^2 two vector fields

$$X_1(x, y) = f_1(y) \partial_x, \quad X_2(x, y) = f_2(x) \partial_y \quad (9.1.5)$$

with arbitrary smooth functions $f_1, f_2 \in C^\infty(\mathbb{R})$. Both of them are complete and have the flows

$$F_t^1(x, y) = (x + t f_1(y), y), \quad F_t^2(x, y) = (x, y + t f_2(x)). \quad (9.1.6)$$

Choosing $f_1 = f_2 = s^2/2$, the sum $X_1 + X_2$ restricted to the diagonal $\Delta \subset \mathbb{R}^2$ defines a vector field on Δ . Using the global coordinate x on Δ , we have $X = x^2 \partial_x$, which we have just shown to be not complete. Hence $X_1 + X_2$ is not complete.

We similarly show that commutators of complete vector fields also need not be complete. Choose $f_1(s) = s$ and $f_2(s) = s^2$. Then

$$[X_1, X_2](x) = -2xy \partial_x + y^2 \partial_y. \quad (9.1.7)$$

This vector field is tangent to the y -axis $x = 0$, and using the global coordinate on it, it restricts to $X = y^2 \partial_y$, which again is not complete.

Example 9.1.6 (Linear vector fields). Let V be a finite-dimensional real vector space. Vector fields on V can be naturally represented as smooth maps $X : V \rightarrow V$, so the flow equation reads $\dot{v} = X(v(t))$. A vector field is called linear if the corresponding mapping is linear, i.e. $X(v) = Av$ for some $A \in \text{End}(V)$. In this case, X is said to be generated by A . As is well known, the solution with initial condition v_0 is $v(t) = e^{tA}v_0$ using the matrix exponential. Thus the flow is given by the “one-parameter group” of automorphisms $F_t = e^{tA} \in \text{GL}(V)$.

Theorem 9.1.7 (Fundamental Theorem on Flows [7, Thm. 9.12]). *Let X be a smooth vector field on a smooth manifold M . There is a unique maximal flow $F : \mathcal{D} \rightarrow M$ whose infinitesimal generator is X , i.e. $\partial_t|_{t=0} F = X$. This flow has the following properties:*

- (a) *For each $p \in M$, the curve F_p is the unique maximal integral curve of X starting at p ;*

(b) If $s \in \mathcal{D}_p$, then $\mathcal{D}_{F(p,s)} = D_p - s$.

Theorem 9.1.8. For any smooth vector field $X \in \mathfrak{X}(M)$, its maximal flow F exists, is unique and:

1. $\mathcal{D}_{F_t(p)} = \mathcal{D}_p - t$ (subtraction of t means a shift of the set on the real line);
2. If t is such that $\mathcal{D}_t = M$, then $F_t : M \rightarrow M$ is a diffeomorphism.

Proof. The (local) existence of an integral curve at p follows from regular existence theorems for ODE's on \mathbb{R}^n . Given two integral curves starting at p , the set of points where they coincide must be closed by the Hausdorff property, and also open by the uniqueness of local solutions to ODE's in \mathbb{R}^n . This means that they must coincide on the entire domain on which they're both defined.

For each $p \in M$, let \mathcal{D}_p be the union of all open intervals in \mathbb{R} containing 0 on which an integral curve starting at p is defined. By the uniqueness that we just shows, this obviously defines a unique maximal integral curve at p .

Part (a) is a simple exercise. Part (b) is a bit more involved but essentially follows from the regular ODE theorem stating that the unique solution depends smoothly on the initial condition in some neighborhood of p . This can be used to show that $\mathcal{D} = \bigcup_p \{p\} \times \mathcal{D}_p$ is open in $M \times \mathbb{R}$ (see [7, Thm. 9.12] for details). Then each set \mathcal{D}_t is open in m and the flow F_t maps \mathcal{D}_t to \mathcal{D}_{-t} . By uniqueness we then see that $F_t \circ F_{-t}$ and $F_{-t} \circ F_t$ are identity maps on \mathcal{D}_t and \mathcal{D}_{-t} , respectively, and thus diffeomorphisms. ■

Corollary 9.1.9 (Naturality of Flows). Let $X \in \mathfrak{X}(M)$, let its flow be F^X and $f \in C^\infty(M, N)$. Then if $Y = f_* X$ and F^Y is the flow of Y , then

$$F_t^Y \circ f = f \circ F_t^X. \quad (9.1.8)$$

Proof. By definition of push-forwards of tangent vectors, if $\gamma_p(t) = F_t^X(p)$ is an integral curve of X , then $f \circ \gamma_p(t)$ is an integral curve of Y :

$$f \circ \gamma_p(t) = F_t^{f_* X}(f(p)), \quad (9.1.9)$$

which is equivalent to the claim. ■

Theorem 9.1.10 (Uniform Time Lemma). Let $X \in \mathfrak{X}(M)$ with flow domain \mathcal{D} . If $\exists \epsilon > 0$ such that $M \times I_\epsilon \subset \mathcal{D}$, then X is complete.

Proof. Suppose X is not complete. Then there exists $p \in M$ such that $\mathcal{D}_p \neq \mathbb{R}$. Then \mathcal{D}_p is bounded by some $b \in \mathbb{R}$. If we choose $t_0 \in (b - \epsilon/3, b)$, then by definition of the flow domain

$F_{\epsilon/2}(F_{t_0}(p))$ exists, but at the same time equals $F_{t_0+\epsilon/2}(p)$, whose existence contradicts the definition of b . ■

Corollary 9.1.11. *If $\text{supp}(X)$ is compact in M , then X is complete. If M is compact, then all vector fields on it are complete.*

Proof. Let $K = \text{supp}(X)$. Every $p \in K$ we have the flow F^X defined on some $U_p \times I_{\epsilon_p}$. Therefore we can cover K by a finite collection of such U_{p_i} 's. Let $\epsilon = \min_k(\epsilon_{p_k})$. Then F^X is well-defined on $K \times I_\epsilon$. Outside of K , the flow is $F_t|_{M \setminus K} = \text{id}$. So the flow is defined on all of M for $t \in I_\epsilon$ and is complete by the Uniform Time Lemma. ■

Conversely, as we will see in the Escape Lemma 9.2.3, if a vector field is not complete, there are integral curves that escape any compact domain in finite time.

9.2 Applications of flows

Definition 9.2.1 (Invariant vector fields on Lie groups). Let G be a Lie group. A vector field $X \in \mathfrak{X}(G)$ is called left-invariant (resp. right-invariant) if for all $g \in G$, if the evaluation map $G \mapsto TG$ given by $g \mapsto X_g$ is equivariant with respect to the left (resp. right) actions of G on itself and on TG . In other words, one of the following formulas must hold for all $h, g \in G$:

$$(L_g)_*X_h = X_{gh}, \quad \text{or} \quad (R_g)_*X_h = X_{hg}. \quad (9.2.1)$$

We will denote by $\mathfrak{X}_L(G)$ the linear space of all left-invariant vector fields on G , and by $\mathfrak{X}_R(G)$ the right-invariant analog.

Note that every invariant vector field is uniquely determined by its value at the identity and is automatically smooth because the left/right action of G on itself is smooth.

Theorem 9.2.2. *All invariant vector fields on Lie groups are complete.*

Proof. Recall that for a smooth map f , the image $f \circ \gamma$ of an integral curve of a vector field X is an integral curve of the transformed field f_*X .

Now let F be the flow of a left-invariant field $X \in \mathfrak{X}(G)$, defined on its flow domain \mathcal{D} . Thus for any $g \in G$, $L_g \circ F_e$ is an integral curve of X starting at g and therefore equal to F_g . This means that we can apply the uniform time lemma and conclude completeness. ■

Lemma 9.2.3 (Escape Lemma). *If an integral curve γ of a vector field X cannot be extended to $t > b$, then $\gamma([t_0, b])$ is not contained in any compact subset of M for any t_0 in the domain of γ .*

Proof. Assume $b < \infty$ and suppose $\gamma([t_0, b))$ is contained in a compact set K . Then if $t_i \nearrow b$, by compactness we have the convergence $\gamma(t_i) \rightarrow q$. Then there is a neighborhood U of q and an $\epsilon > 0$ such that the flow F is defined on $U \times I_\epsilon$. Pick i large enough so that $t_i > b - \epsilon$ and $\gamma(t_i) \in U$. Then

$$\tilde{\gamma}(t) = \begin{cases} \gamma(t), & t \in [t_0, b), \\ F_{t-t_i} \circ F_{t_i}(p), & t \in (t_i - \epsilon, t_i + \epsilon) \end{cases} \quad (9.2.2)$$

is a well-defined integral curve because $t - t_i < t_i + \epsilon - t_i = \epsilon$, so F_{t-t_i} exists by assumption. This contradicts the assumption that γ cannot be extended. ■

Now we are finally ready to prove Ehresmann's Fibration Theorem 7.1.8, which states that every proper surjective submersion is a locally trivial fibration.

Proof of Ehresmann's Fibration Theorem 7.1.8. First we note that since the statement is local, it suffices to prove it for $M = \mathbb{R}^n$. Locally (both in E and M), all submersions are trivial fibrations. The idea is to use flows of vector fields along the fibers of the submersion and a POU to glue these trivial pieces together into a locally trivial fibration.

We will map the fibers $\pi^{-1}(\mathbb{R}^n)$ to the single fiber $\pi^{-1}(0)$ in a diffeomorphic way. This is done by moving the fibers along integral curves of some set of vector fields, in such a way that the flow takes each fiber to $\pi^{-1}(0)$. These vector fields are chosen to be π -related to some local sections of TM , so as to satisfy all the requirements.

Here are the details. Due to the fact that π is a submersion, there exist vector fields $X_i \in \mathfrak{X}(E)$ that are π -related to the coordinate vector fields ∂_i on $M = \mathbb{R}^n$, $i = 1, \dots, n$. Indeed, this is obvious locally (both in E and M), but a collection of such local vector fields X_i^α on $U_\alpha \subset E$ can be used to define a global vector field $X_i = \sum \chi_\alpha X_i^\alpha$ using a POU χ_α on E .

We claim that all of the vector fields X_i are complete on E . Suppose $e \in E$ and the flow $F^i(t, e)$ of X^i is defined for t in some finite open interval in \mathbb{R} . Since the flow of ∂_i in the base is given by translations at a constant velocity, the image of this open interval under $\pi \circ F^i(_, e)$ must be contained inside some compact set $K \subset M$. Hence the image of a finite open interval under $F^i(_, e)$ must be contained in $\pi^{-1}(K)$, which is compact because π is proper. If $F^i(_, e)$ were not defined on all of t , its values would have to escape any compact subset of E , leading to a contradiction. Hence X_i is complete.

Now let F_t^i be the flow of X_i and define $G : \pi^{-1}(\mathbb{R}^n) \rightarrow \mathbb{R}^n \times \pi^{-1}(0)$ by

$$G(z) = \left(\pi(z), F_{-t_n}^n \circ \dots \circ F_{-t_1}^1(z) \right), \quad \text{where } \pi(z) = (t_1, \dots, t_n). \quad (9.2.3)$$

Clearly G is a diffeomorphism and a local trivialization for π . ■

Definition 9.2.4 (Flow-outs). Let $S \lhd M$ as a subset and pick a vector field $X \in \mathfrak{X}(M)$ that is nowhere parallel to S , i.e. $X(p) \notin TS$ for all $p \in S$. If \mathcal{D} is the flow domain of X , then $\Phi = F|_{\mathcal{O}}$ with $\mathcal{O} = (S \times \mathbb{R}) \cap \mathcal{D}$ is called a flow-out of S .

Theorem 9.2.5 (Flow-out lemma). *With the notation of the above definition,*

1. *Φ is an immersion $\Phi(\mathcal{O}_\delta) \lhd M$ as long as $\delta : S \rightarrow \mathbb{R}_{>0}$ is chosen so that when restricted to the set $\mathcal{O}_\delta = \{(p, t) \in \mathcal{O} \mid |t| < \delta(p)\}$, $\Phi(\mathcal{O}_\delta)$ is injective. Moreover, such δ always exists;*
2. *$\Phi_*(\partial_t) = X$ on \mathcal{O} ;*
3. *If $\text{codim } S = 1$, then $\Phi(\mathcal{O}_\delta) \lhd M$.*

Proof. Exercise. The last part follows because the dimension of $\Phi(\mathcal{O}_\delta)$ coincides with the dimension of M , so injectivity and local diffeo implies embedding. ■

Let ∂M be non-empty and let $\{U_\alpha\}$ be a family of boundary charts that cover ∂M . On each of them, define the local coordinate field $X_\alpha = \partial_n$ that points inward from the boundary. Further, let χ_α be a subordinate POU. Then

$$N = \sum_\alpha \chi_\alpha X_\alpha \quad (9.2.4)$$

is a vector field defined on a neighborhood U of ∂M . In the local boundary coordinates, $N^n = \sum \chi_\alpha X_\alpha^n > 0$, therefore N is strictly *inward*.

Since $\text{codim } \partial M = 1$, we only need to find a δ to construct a flow-out of ∂M . With δ , we have an embedding $\Phi : \mathcal{O}_\delta \rightarrow M$. Set

$$\psi : [0, 1] \times \partial M \rightarrow \mathcal{O}_\delta, \quad (p, t) \mapsto (p, t\delta(p)). \quad (9.2.5)$$

Then ψ is a diffeo and $\Phi \circ \psi : [0, 1] \times \partial M \rightarrow \Phi(\mathcal{O}_\delta)$ is also a diffeo, so

$$\Phi(\mathcal{O}_\delta) \cong [0, 1] \times \partial M. \quad (9.2.6)$$

Definition 9.2.6 (Collar). The embedded submanifold $\Phi(\mathcal{O}_\delta)$ of a manifold M with a boundary constructed above is called a collar of M .

Corollary 9.2.7. *By varying t , we find from the above construction that M is homotopy equivalent to $\overset{\circ}{M}$.*

Now, if we have two manifolds M and N with diffeomorphic boundaries, so that

$$h : \partial M \rightarrow \partial N - \text{diffeo}, \quad (9.2.7)$$

then we can introduce two collar embeddings (which are also diffeo onto their images)

$$\alpha : [0, 1] \times \partial M \rightarrow V \subset M, \beta : [0, 1] \times \partial N \rightarrow W \subset N, \quad (9.2.8)$$

and unite them by

$$\Phi : V \sqcup W \rightarrow \partial M \times (-1, 1), \quad \Phi(x) = \begin{cases} (p, -t), & x = \alpha(p, t) \in V, \\ (h(q), t), & x = \beta(q, t) \in W. \end{cases} \quad (9.2.9)$$

Clearly Φ is a topological embedding with a closed image (it's surjective), therefore Φ is a closed mapping.

Now let $\pi : M \sqcup N \rightarrow X$ be the quotient map on whose fibers Φ is constant. Then Φ descends to $\tilde{\Phi} : X \rightarrow [-1, 1] \times \partial M$. Φ is continuous and bijective, $\tilde{\Phi}(K) = \Phi(\pi^{-1}(K))$ is closed for any compact K (because Φ is closed and π is continuous), so $\tilde{\Phi}$ is closed too. Therefore $\tilde{\Phi}$ is a homeomorphism and therefore a diffeomorphism, and X is a quotient manifold. Alternatively, this follows from the so-called *closed graph theorem* that we won't cover here.

Definition 9.2.8 (Gluing of manifolds). The construction described above is called the gluing of two manifolds with diffeomorphic boundaries.

Definition 9.2.9 (Connected sum). Let $M_i, i = 1, 2$ be two connected manifolds without boundaries and $M'_i = M_i \setminus U_i$, where U_i is a coordinate ball. Then $\partial M'_i \cong \mathbb{S}^{m-1}$ and let the diffeomorphisms here be h_i . Then we define the connected sum as the gluing along $h_1 \circ h_2^{-1}$:

$$M_1 \# M_2 := M'_1 \cup_{h_1 \circ h_2^{-1}} M'_2. \quad (9.2.10)$$

Example 9.2.10. $D(M) = M \cup_{\partial M} M$ is called the *double* of a manifold with a boundary.

Example 9.2.11. If $X \parallel \partial D$, $D \subset M$, then F^X leaves D invariant.

Corollary 9.2.12. If $X \parallel \partial M$, then the theorem about the existence of flows applies to X .

Proof. Extend X to the double $D(M)$, apply the flow theorem there, then restrict back to M . ■

Example 9.2.13. So far we've only considered time-independent vector fields. In general, we can have a time-dependent vector field $X : M \times J \rightarrow TM$, $J \subset \mathbb{R}$. All of the results about flows still hold almost unchanged, only now the flow map ("evolution operator") depends on two time parameters: $F^X(p, t_1, t_2)$.

Theorem 9.2.14 (Rectification/Straightening Theorem). *If $X \in \mathfrak{X}(M)$ and $X(p) \neq 0$, then there exists a local chart (U, φ) at p such that its slice $S = \{p \in U \mid \varphi^1(p) = 0\}$ is*

diffeomorphic to \mathbb{S}^{m-1} , $X(p) \notin T_p S$ and $X = \partial_1$ on U .

Proof. Since $S \lhd M$, it has a slice neighborhood. We can parametrize S by an embedding map $x : \Omega \rightarrow S$ with $\Omega \subset \mathbb{R}^{m-1}$. We can take U such that $X|_U \cap TS = \emptyset$.

Now we can use a flow-out along X so that the flow $F_t(x(s))$ parametrizes $F(x(\Omega)) \lhd M$. This is a diffeo $O_\delta \rightarrow F(x(\Omega)) = x(\Omega)_\delta$ and hence has a smooth inverse. The inverse then gives a local chart $\varphi = \psi^{-1}$, where $\psi(s, t) = F_t(x(s))$.

Finally, $X = \partial_t = \partial_1$ in this local chart. ■

Definition 9.2.15 (Complete solution). A complete solution of a vector field $X \in \mathfrak{X}(M^m)$ is a local function $\Psi : U \times I_b \rightarrow \mathbb{R}^m$ such that $\Psi^{-1}(c)$ for every c in the image of Ψ is an integral curve of X at a $p \in U$. The components Ψ^i of this map are called a *complete set of integrals of motion* for X .

Proposition 9.2.16. *Local complete solutions exists for all vector fields.*

Proof. First rectify the vector field, then the coordinates of the point of intersection of each integral curve with the slice S will be the (constant) value of Ψ on this whole curve. ■

9.3 Vector fields as differential operators

Recall that Lie derivatives of functions satisfy $\mathcal{L}_{F_* X} F_* f = F_*(\mathcal{L}_X f)$ for diffeomorphisms F . Also, Lie derivatives are local: for any open set U

$$\mathcal{L}_{X|_U} (f|_U) = (\mathcal{L}_X f)|_U. \quad (9.3.1)$$

Definition 9.3.1 (Derivations). If A is an algebra over a field K , then a derivation on A is a K -linear map $D : A \rightarrow A$ that satisfies the Leibniz rule

$$D(a \cdot b) = D(a) \cdot b + a \cdot D(b), \quad a, b \in A. \quad (9.3.2)$$

It is easy to check that the vector space of all derivations on A , denoted by $\text{Der}(A)$, is a Lie algebra under the commutator $[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1$. An algebra with a distinguished derivation is called a *differential algebra*.

Clearly any Lie derivative \mathcal{L}_X , $X \in \mathfrak{X}(M)$, is a derivation on the algebra of functions $C^\infty(M)$, because it is a directional derivative in local coordinates.

Moreover, any derivation D on $C^\infty(M)$ satisfies $D(1) = 0$, because

$$D(1) = D(1^2) = D(1) \cdot 1 + 1 \cdot D(1) = 2 \cdot D(1). \quad (9.3.3)$$

Finally, the set of all Lie derivatives $\{\mathcal{L}_X\}_{X \in \mathfrak{X}(M)}$ is a $C^\infty(M)$ -module:

$$f \cdot \mathcal{L}_X = \mathcal{L}_{f \cdot X}. \quad (9.3.4)$$

Proposition 9.3.2. *The space of Lie derivatives \mathcal{L}_X is isomorphic to $\mathfrak{X}(M)$ as $C^\infty(M)$ -modules. In particular, $\mathcal{L}_X = 0$ iff $X = 0$.*

Proof. The first part is obvious. If $d_p f(X_p) = 0$ for all $f \in C^\infty(M)$, then $\alpha_p(X_p) = 0$ for all $\alpha_p \in T_p^* M$. But then $X_p = 0$. ■

Proposition 9.3.3. *The space of all \mathbb{R} -linear derivations on the algebra $C^\infty(M)$ is isomorphic to $\mathfrak{X}(M)$ as a vector space. In particular, every derivation θ coincides with \mathcal{L}_X for a unique $X \in \mathfrak{X}(M)$.*

Proof. Given θ , we construct the vector field X . First of all, θ is local, i.e. if h vanishes on a neighborhood $V \subset M$ of $p \in V$, then $\theta(h)(p) = 0$. Indeed, if χ is a bump function equal to 1 around p and zero outside of V , then $h = (1 - \chi)h$ and

$$\theta(h)(p) = \theta(\chi h + (1 - \chi)h)(p) = (\theta(1 - \chi)) \cdot h(p) + \theta(h) \cdot (1 - \chi(p)) = 0. \quad (9.3.5)$$

Pick a local chart at p , $\varphi : U \rightarrow \hat{U} \subset \mathbb{R}^m$. Let $x \in \hat{U}$, $a = \varphi(p)$, $f \in C^\infty(M)$, then

$$(\varphi_* f)(x) = (\varphi_* f)(a) + \sum_{j=1}^m (x^j - a^j) \int_0^1 \underbrace{(\varphi_* f)_j}_{\partial(\varphi_* f)/\partial x^j} [a + t(x - a)] dt \quad (9.3.6)$$

on some neighborhood of a .

Therefore

$$f(u) = f(p) + \sum_{i=1}^m (\varphi^i(u) - a^i) g_i(u) \quad (9.3.7)$$

for some set of local functions, namely $g_i(p) = \frac{\partial(\varphi_* f)}{\partial x^i}(a)$. Thus

$$(\theta f)(p) = \sum_i g_i(p) \theta(\varphi^i)(p) = \sum_i \partial_{x^i}(\varphi_* f)(a) \theta \varphi^i(p), \quad (9.3.8)$$

and this formula is chart-independent! Now we can define a smooth X by its local representation

$$X^i(u) := \theta \varphi^i(u). \quad (9.3.9)$$

It is easy to check that X is independent of the choice of φ , which means that it is actually a globally well-defined vector field on M . For $f \in C^\infty(M)$, the local representative in $x = \varphi(u) \in \hat{U}$ of $\mathcal{L}_X f$ is

$$(f \circ \varphi^{-1})_{*x} \circ \hat{X}(x) = \sum_i \partial_{x^i} (f \circ \varphi^{-1})(x) \cdot \theta \varphi^i(u) = \theta f(u), \quad (9.3.10)$$

hence $\theta = \mathcal{L}_X$.

Uniqueness follows from the statement that $\mathcal{L}_X = 0$ iff $X = 0$. ■

Corollary 9.3.4. *If x^i is a local coordinate system, then the local vector fields $\partial/\partial x^i$ form a basis of derivations at p , therefore*

$$(\mathcal{L}_X f)(p) = \sum_i \partial_{x^i} (f \circ \varphi^{-1})(\varphi(p)) \cdot \underbrace{\mathcal{L}_X(\varphi^i)(p)}_{X^i} = \sum_i X^i \partial_{x^i} (\varphi_* f) = \sum_i X^i \partial_i f. \quad (9.3.11)$$

Proposition 9.3.5. *If $X, Y \in \mathfrak{X}(M)$, then $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X$ is a derivation on $C^\infty(M)$.*

Proof. This is a very general property that the commutator of two derivations is always a derivation:

$$\begin{aligned} [D_1, D_2](fg) &= D_1((D_2 f)g + f(D_2 g)) - D_2((D_1 f)g + f(D_1 g)) = \\ &= ([D_1, D_2]f)g + f([D_1, D_2]g). \end{aligned} \quad (9.3.12) \quad \blacksquare$$

This proposition allows us to define the commutator of vector fields.

Definition 9.3.6 (Lie derivative/bracket of vector fields). If $X, Y \in \mathfrak{X}(M)$, then the Lie derivative of Y along X is denoted by

$$\mathcal{L}_X Y \equiv [X, Y] \quad (9.3.13)$$

and is defined as the unique vector field such that

$$\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]. \quad (9.3.14)$$

With this bracket, $\mathfrak{X}(M)$ becomes a Lie algebra that is isomorphic to the Lie algebra of derivations on $C^\infty(M)$ (i.e. Lie derivatives). In local coordinates (using Einstein summation),

$$[X, Y]^j = X^i \partial_i Y^j - Y^i \partial_i X^j. \quad (9.3.15)$$

Proposition 9.3.7. 1. $[X, Y]$ is local and natural w.r.t. diffeomorphisms, i.e.

$$F_*[X, Y] = [F_*X, F_*Y]; \quad (9.3.16)$$

2. \mathcal{L}_X is a derivation on the Lie algebra $\mathfrak{X}(M)$, i.e. satisfies the Jacobi identity

$$\mathcal{L}_X[Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]; \quad (9.3.17)$$

3. \mathcal{L}_X is also a derivation on $C^\infty(M) \otimes \mathfrak{X}(M)$ i.e. it satisfies the mixed Leibniz rule

$$\mathcal{L}_X(f \otimes Y) = \mathcal{L}_X(f \cdot Y) = (\mathcal{L}_X f) \cdot Y + f \cdot \mathcal{L}_X Y. \quad (9.3.18)$$

Equivalently,

$$\mathcal{L}_{fX} Y = f \cdot \mathcal{L}_X Y + (\mathcal{L}_X f) \cdot Y. \quad (9.3.19)$$

Proof. 1. To prove any of these properties, we act with both sides on a function f as a Lie derivative. Then the equality of the two expressions as derivations implies the equality as vector fields by Proposition 9.3.3. Therefore this is just a direct computation using the push-forward property for Lie derivatives acting on functions.

2. General property of derivations that is easily checked by expanding both sides.
3. Check

$$\begin{aligned} [X, fY]g &= \mathcal{L}_X(\mathcal{L}_{fY} g) - \mathcal{L}_{fY}\mathcal{L}_X g = \mathcal{L}_X(f\mathcal{L}_Y g) - f\mathcal{L}_Y\mathcal{L}_X g = \\ &= (\mathcal{L}_X f)\mathcal{L}_Y g + f\mathcal{L}_X\mathcal{L}_Y g - f\mathcal{L}_Y\mathcal{L}_X g = ((\mathcal{L}_X f)Y + f[X, Y])g. \end{aligned} \quad (9.3.20)$$

■

Remark 9.3.8 (Lie derivatives are not covariant derivatives). Property (9.3.19) shows that $\mathcal{L}_X Y$ doesn't behave like a tensor with respect to X (or Y), since it is not $C^\infty(M)$ -linear in it. This is why Lie derivatives are not a special case of *covariant* derivatives.

A covariant derivative (of vector fields) is a map $\nabla : \mathfrak{X}(M) \otimes \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$ denoted by $X, Y \mapsto \nabla_X Y$ such that it satisfies:

- $C^\infty(M)$ -linearity in X : $\nabla_{fX} Y = f \cdot \nabla_X Y$ for $f \in C^\infty(M)$;
- \mathbb{R} -linearity in Y : $\nabla_X(\alpha Y + Z) = \alpha \nabla_X Y + \nabla_X Z$ for $\alpha \in \mathbb{R}$;
- Leibniz rule: $\nabla_X(f \cdot Y) = (\mathcal{L}_X f) \cdot Y + f \cdot \nabla_X Y$.

Therefore any covariant derivative $\nabla_X Y$ has to be a tensor in X , unlike Lie derivatives. In other words, a covariant derivative can be thought of as a map $\nabla : \mathfrak{X}(M) \rightarrow \mathfrak{X}^*(M) \otimes \mathfrak{X}(M) = \Gamma_1^1(M)$, $Y \mapsto \nabla Y \in \Gamma_1^1(M)$.

All of this allows us to extend the Lie derivative to tensor fields of any rank by the following theorem.

Definition 9.3.9 (Tensor differential operators). A differential operator on the tensor algebra $\Gamma(M)$ is a collection of maps $\mathbb{D}_s^r(U) : \Gamma_s^r(U) \rightarrow \Gamma_s^r(U)$ for all $r, s \geq 0$ and each open $U \subset M$ (we denote them all by \mathbb{D}) such that:

1. \mathbb{D} is a tensor derivation (w.r.t. \otimes);
2. \mathbb{D} is local, i.e. for $U \subset V$ and $t \in \Gamma_s^r(V)$, $\mathbb{D}(t)|_U = \mathbb{D}(t|_U)$;
3. $\mathbb{D}\delta = 0$, where $\delta \in \Gamma_1^1(M)$ is the Kronecker tensor $\delta = \text{id}_{TM}$.

Note that we don't require $\mathbb{D} \circ \varphi_* = \varphi_* \circ \mathbb{D}$ because this will not hold for covariant derivatives, which we would still like to consider differential operators.

Theorem 9.3.10 (Willmore). *A collection of local tensor derivations E_U on $C^\infty(U)$ and F_U on $\mathfrak{X}(U)$ (which agree with each other via the mixed rank Leibniz rule) for each $U \subset M$ determines a unique differential operator \mathbb{D} on $\Gamma(M)$ that coincides with E_U on C^∞ and with F_U on \mathfrak{X} .*

Proof. Suppose we know that such a \mathbb{D} exists. In a local frame $\{e_i\}$ and dual frame $\{\alpha^i\}$, we can write

$$t(u) = \underbrace{t_{j_1 \dots j_s}^{i_1 \dots i_r}(u)}_{\in C^\infty(U)} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes \alpha^{j_1} \otimes \dots \otimes \alpha^{j_s}(u). \quad (9.3.21)$$

By using the tensorial Leibniz rule for \mathbb{D} , we express $\mathbb{D}t(u)$ only in terms of $E_U(t_{j_1 \dots j_s}^{i_1 \dots i_r})$ and $F_U(e_{i_k})$.

Namely, since $\mathbb{D}\delta = 0$, and $\delta = e_j \otimes \alpha^j$ (Einstein summation), we have

$$0 = \mathbb{D}(e_j \otimes \alpha^j) = (\mathbb{D}e_j) \otimes \alpha^j + e_j \otimes (\mathbb{D}\alpha^j). \quad (9.3.22)$$

Since $\delta_j^i = \alpha^i(e_j)$, if we contract the formula above with e_i , we get

$$0 = F_U e_i + \langle \mathbb{D}\alpha^j, e_i \rangle e_j, \quad (9.3.23)$$

which completely determines $\mathbb{D}\alpha^j$. Therefore the action of \mathbb{D} on \mathfrak{X}^* is uniquely determined, and the Leibniz rule applied to 9.3.22 gives a formula for \mathbb{D} in terms of E_U and F_U on the whole tensor algebra.

The formulas we have derived prove the uniqueness of \mathbb{D} . Its existence follows from the fact that these same formulas define a valid differential operator that satisfies all the requirements of the theorem. ■

Corollary 9.3.11. A byproduct of the theorem above is the general formula for differential operators on the tensor algebra:

$$\begin{aligned} \mathbb{D} \left(t(\alpha^1, \dots, \alpha^r, X_1, \dots, X_s) \right) &= (\mathbb{D}t)(\alpha^1, \dots, \alpha^r, X_1, \dots, X_s) + \\ &+ \sum_{j=1}^r t \left(\alpha^1, \dots, \mathbb{D}\alpha^j, \dots, \alpha^r, X_1, \dots, X_s \right) + \sum_{k=1}^s t \left(\alpha^1, \dots, \alpha^r, X_1, \dots, \mathbb{D}X_k, \dots, X_s \right). \end{aligned} \quad (9.3.24)$$

This formula essentially states that \mathbb{D} “commutes with contractions”.

Definition 9.3.12 (Lie derivatives of tensors). If $X \in \mathfrak{X}(M)$ then the Lie derivative operator \mathcal{L}_X on the tensor algebra $\Gamma(M)$ is defined as the unique tensor differential operator (established by Willmore's theorem) that coincides with the previously constructed Lie derivative on $C^\infty(M)$ and $\mathfrak{X}(M)$.

Proposition 9.3.13 (Naturality of Lie derivatives). For a diffeomorphism F ,

$$\mathcal{L}_{F_*X}(F_*t) = F_*(\mathcal{L}_X t). \quad (9.3.25)$$

Proof. Define an operator \mathbb{D} by $\mathbb{D}t = F^* \mathcal{L}_{F_*t}(F_*t)$. We know that \mathbb{D} coincides with \mathcal{L}_X on C^∞ and $\mathfrak{X}(M)$. It is also easy to check that \mathbb{D} is a tensor differential operator (Exercise). Then by Willmore's theorem it has to coincide with \mathcal{L}_X on the entire tensor algebra. ■

9.4 Geometric approach to Lie derivatives

Let $X \in \mathfrak{X}(M)$ and let (U, a, F) be a flow box of X at $p \in M$. Let $t \in \Gamma_s^r(M)$ and define

$$t_\lambda := F_\lambda^* \left(t \Big|_{U_\lambda} \right) = F_{\lambda^*}^{-1} \left(t \Big|_{U_\lambda} \right) \in \Gamma_s^r(U), \quad (9.4.1)$$

where $F_\lambda : U \rightarrow U_\lambda$ is the flow diffeomorphism. Then with fixed p , the map $\lambda \mapsto t_\lambda(p) \in \mathbb{T}_s^r(T_p M)$ defines a curve in the vector space of tensors at p .

Theorem 9.4.1. $\lambda \mapsto t_\lambda(p)$ defined above is a curve in the vector space $\mathbb{T}_s^r(T_p M)$ passing through $t(p)$ and its derivative evaluates the Lie derivative of t :

$$\mathcal{L}_X t(p) = \frac{d}{d\lambda} \Big|_{\lambda=0} t_\lambda(p). \quad (9.4.2)$$

Proof. We have $t_\lambda(p) = (TF_\lambda^{-1})_s^r t(F_\lambda(p))$, which is clearly smooth.

Define θ_X as an operator on $\Gamma_s^r(M)$ by $\theta_X t(p) = \frac{d}{d\lambda}|_{\lambda=0} t_\lambda(p)$. For $t \in C^\infty(M)$, $t_\lambda(p)$ is smooth, therefore $\theta_X t(p) = (Tt_\bullet)(\partial_\lambda) \in C^\infty(M)$, where the differential T is with respect to the variable λ (represented by \bullet) and ∂_λ is a vector field on the \mathbb{R} . θ_X is a linear derivation because its local representatives are. It is defined locally, therefore it is local. Finally, $\theta_X \delta = \frac{d}{d\lambda}|_{\lambda=0} F_\lambda^* \delta = \frac{d}{d\lambda}|_{\lambda=0} \delta = 0$. This proves that θ_X is a differential operator. It remains to check that $\theta_X = \mathcal{L}_X$ on $C^\infty(M)$ and $\mathfrak{X}(M)$.

For $f \in C^\infty(M)$, $f_\lambda = f \circ F_\lambda$, $\frac{d}{d\lambda}|_{\lambda=0} f_\lambda(p) = df \circ \frac{d}{d\lambda}|_{\lambda=0} F_\lambda = df \circ X = \mathcal{L}_X f$.

Now, note that for $f \in C^\infty(M)$ there exists a $g_\lambda \in C^\infty(M)$ such that $g_0 = \mathcal{L}_X f$ and $f \circ F_\lambda = f + \lambda g_\lambda$. Namely, $\lambda g_\lambda(p) = \int_0^1 \partial_t(f \circ F_{t,\lambda}(p)) dt$. Then for $Y \in \mathfrak{X}(M)$,

$$\begin{aligned} \mathcal{L}_{\theta_X Y} f(p) &= df \circ \frac{d}{d\lambda} \Big|_{\lambda=0} ((F_{-\lambda})_* p \circ Y \circ F_\lambda(p)) = \frac{d}{d\lambda} \Big|_{\lambda=0} ((f \circ F_{-\lambda})_* \circ Y \circ F_\lambda(p)) = \\ &= \frac{d}{d\lambda} \Big|_{\lambda=0} ((f - \lambda g_{-\lambda})_* \circ Y \circ F_\lambda(p)) = \frac{d}{d\lambda} \Big|_{\lambda=0} f_* \circ Y \circ F_\lambda(p) - \left(\frac{d}{d\lambda} \lambda \right) (g_{-\lambda})_* \circ Y \circ F_\lambda(p) \Big|_{\lambda=0} = \\ &= d(\underbrace{f_* \circ Y}_{\mathcal{L}_Y f}) \circ \underbrace{\frac{d}{d\lambda} \Big|_{\lambda=0} F_\lambda(p)}_{X(p)} - (g_0)_* \circ Y \circ F_0(p) = \mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f(p), \quad (9.4.3) \end{aligned}$$

therefore $\theta_X Y = [X, Y]$. By Willmore's theorem, $\theta_X = \mathcal{L}_X$. ■

Corollary 9.4.2. 1. $\mathcal{L}_X t = 0$ iff t is constant along the flow of X : $F_\lambda^* t = t$;

2. For all λ in the flow domain the following equality holds:

$$\frac{d}{d\lambda} F_\lambda^* t = F_\lambda^* \mathcal{L}_X t. \quad (9.4.4)$$

In particular,

$$\mathcal{L}_X t = F_{-\lambda}^* \frac{d}{d\lambda} F_\lambda^* t = \frac{d}{d\lambda} \Big|_{\lambda=0} F_\lambda^* t. \quad (9.4.5)$$

Remark 9.4.3 (Push-forwards are not parallel transports). Since covariant derivatives are in the exact same relationship with parallel transports as Lie derivatives are with push-forwards, it is tempting to think that Lie derivatives are special cases of covariant derivatives. However, we have already seen this to be wrong in Remark 9.3.8. This means that push-forwards by flows of vector fields cannot be interpreted as a kind of parallel transport along the integral curves of those vector fields. This is because a parallel transport has to be unambiguously defined for every curve on M , whereas there are infinitely many vector fields for which a

given curve is integral (and the value of the push-forward will depend on the specific choice of vector field). Transport by push-forwards is often called *Lie transport*. Lie derivatives are just infinitesimal Lie transport.

Remark 9.4.4. We can apply these ideas to PDEs on \mathbb{R}^n :

$$\begin{cases} \partial_t f(x, t) = \sum_i X^i \partial_i f(x, t), \\ f(x, 0) = g(x). \end{cases} \quad (9.4.6)$$

Then if X has a complete flow F_t , the unique solution of the above Cauchy problem is $f(x, t) = g(F_t(x))$, which can be checked by substitution. This is sometimes called the *method of characteristics*. The “characteristics” are the integral curves along which the values of $g(x)$ get propagated.

Proposition 9.4.5. Given two vector fields X, Y with complete flows F_t and G_t , respectively, **TFAE**:

1. $[X, Y] = 0$;
2. $F_t^*Y = Y$ (or $G_t^*X = X$);
3. $F_t \circ G_s = G_s \circ F_t$ for all t, s .

In particular, the flow of $X + Y$ for commuting X and Y is $F_t \circ G_t$.

Proof. $1 \Rightarrow 2$. $\frac{d}{dt}F_t^*Y = F_t^*\mathcal{L}_XY = 0$, therefore F_t^*Y is constant and we know that $F_0^*Y = Y$, so $F_t^*Y = Y$.

$2 \Rightarrow 3$. Consider the curve $\gamma(t) = G_s \circ F_t(p)$. It starts at $\gamma(0) = G_s(p)$ and its velocity is

$$\dot{\gamma}(t) = \frac{d}{dt}G_s \circ F_t(p) = G_{s*} \circ X \circ F_t \stackrel{(2)}{=} X \circ \gamma(t). \quad (9.4.7)$$

So $\gamma(t)$ is an integral curve of X going through p . Therefore, by uniqueness (and definition of the flow of X),

$$\gamma(t) = F_t \circ G_s(p). \quad (9.4.8)$$

This shows that the flows commute. Finally, we check

$$\frac{d}{dt}F_t \circ G_t(p) = X(F_t(G_t(p))) + \underbrace{F_{t*}Y}_{Y}(G_t(p)) = (X + Y)(F_t \circ G_t(p)). \quad (9.4.9)$$

$$2 \Rightarrow 1. \mathcal{L}_XY = F_{-t}^* \frac{d}{dt}F_t^*Y = F_{-t}^* \frac{d}{dt}Y = 0.$$

$3 \Rightarrow 2$. $F_{t*}Y = F_{t*} \frac{d}{ds}G_s = F_{t*}G_{s*}Y = (F_t \circ G_s)_*Y = (G_s \circ F_t)_*Y = G_{s*}F_{t*}Y$. But the only vector field invariant under G_{s*} is Y , therefore $F_{t*}Y = Y$. ■

Remark 9.4.6 (Taylor formula on manifolds). Successive application of the vector fields X_1, \dots, X_r to a function $f \in C^\infty(M)$ yields

$$X_r \cdots X_1 \cdot f = \frac{d}{dt_1} \Big|_{t_1=0} \cdots \frac{d}{dt_r} \Big|_{t_r=0} f \circ F_{t_1}^{X_1} \circ \cdots \circ F_{t_r}^{X_r}. \quad (9.4.10)$$

In particular, the Lie bracket can be expressed as

$$[X, Y]f = \frac{d}{dt} \Big|_{t=0} \frac{d}{ds} \Big|_{s=0} \left(f \circ F_s^Y \circ F_t^X - f \circ F_s^X \circ F_t^Y \right). \quad (9.4.11)$$

Furthermore, for any $m \in M$, the Taylor expansion of the smooth function $t \mapsto f \circ F_t^X(m)$ at $t = 0$ yields the following *Taylor formula for manifolds*:

$$f \circ F_t^X(m) = \sum_{k=1}^n \frac{t^k}{k!} (X^k f)(m) + O(t^{n+1}), \quad t \in \mathcal{D}_m^X, \quad (9.4.12)$$

where $O(t^{n+1})$ is a smooth function on \mathcal{D}_m^X such that $O(t^{n+1})/t^{n+1}$ is bounded. Repeated application of this formula yields the iterated Taylor formula for manifolds

$$f \circ F_t^X \circ F_s^Y(m) = \sum_{k,l=1}^n \frac{t^k s^l}{k! l!} (Y^l X^k f)(m) + O(t^{n+1}, s^{n+1}), \quad (9.4.13)$$

which easily generalizes to any number of vector fields.

Definition 9.4.7 (Holonomic frame). A local k -frame $\{e_i\}_{i=1}^k$ is called holonomic if $[e_i, e_j] = 0$ for all i, j .

Theorem 9.4.8 (Rectification of frames). Let $\{e_j\}_{j=1}^k$ with $k \leq m = \dim M$ be a holonomic local k -frame on $W \subset M$. Assume that $e_j(p) \notin T_p S$ for all $p \in S$, where $S \subset M$ is a submanifold with $\text{codim } S = k$. Then there exists a local chart of coordinates (s^1, \dots, s^m) in which $S = \{s^1 = \cdots = s^k = 0\}$ and for $j \leq k$ we have $\partial_j = e_j$.

Proof. Parametrize S as usual by $x : \Omega \rightarrow S$ with $\Omega \subset \mathbb{R}^{m-k}$, define a chart $\varphi = \psi^{-1}$, where

$$\psi(s^1, \dots, s^k, s^{k+1}, \dots, s^m) = F_{s_1}^{e_1} \circ \cdots \circ F_{s_k}^{e_k}(x(s^{k+1}, \dots, s^m)), \quad (9.4.14)$$

and F^{e_i} is the flow of e_i . The inverse ψ^{-1} exists locally because the differential at $(0, \dots, 0)$

is identity. Finally we use commutation to compute the partials

$$\begin{aligned}\psi_*(\partial_{s^i}) \cdot f &= \partial_i(f \circ \psi) = \partial_i f(F_{s_1}^{e_1} \cdots F_{s_k}^{e_k} x(0, \dots, 0, s^{k+1}, \dots)) = \\ &= \partial_i f(F_{s_i}^{e_i} \circ \cdots) = \partial_i F_{s_i}^{e_i}(F \cdots F(0, \dots, 0, s^{k+1}, \dots))f = e_i \cdot f = \mathcal{L}_{e_i} f,\end{aligned}\quad (9.4.15)$$

thus $\partial_i = e_i$. ■

Corollary 9.4.9. *A helpful way to reformulate this theorem for $k = m$ is: a given local frame is a coordinate frame for some local coordinates if and only if it is holonomic.*

The analog of this Corollary for $k < m$ is a famous theorem that deserves its own section.

9.5 Frobenius theorem I

Recall that a distribution D on a manifold is an assignment of a subspace $D_p \subset T_p M$ for every $p \in M$. If the dimension of D_p is constant and equal to k , then D is said to be of rank k . D is called a *smooth distribution* if it is a subbundle of TM .

Definition 9.5.1 (Integral manifolds). Let $D \subset TM$ be a smooth distribution (i.e. a subbundle of TM). An immersed submanifold $N < M$ is called an integral manifold of D if $T_p N = D_p$ for all $p \in N$.

Definition 9.5.2 (Integrable distributions). A smooth distributions $D \subset TM$ is called integrable if every $p \in M$ is contained in a (local) integral manifold of D .

Definition 9.5.3 (Involutive distributions). For a smooth distribution $D \subset TM$, denote by $\Gamma(D) \subset \mathfrak{X}(M)$ the space of D -valued vector fields. D is called involutive if $\Gamma(D)$ is a Lie subalgebra of $\mathfrak{X}(M)$, i.e. $[X, Y] \in \Gamma(D)$ for all $X, Y \in \Gamma(D)$.

Theorem 9.5.4 (Frobenius Theorem [7, Thm. 19.12]). *A smooth distribution D is involutive iff it is integrable.*

Proof. The backwards direction is trivial: if D is integrable, then $D = TN$ locally, which implies involutivity.

For the forward direction, it suffices to prove that D is locally spanned by a set of linearly independent commuting fields. Then by the rectification theorem they are coordinate vector fields for a slice chart of an integral manifold.

Because of the locality of the statement, it also suffices to prove the theorem for $M = U \subset \mathbb{R}^m$. At $p \in M$, we can find a system of Cartesian coordinates such that D_p is complementary to $\text{Span}(\partial_{k+1}, \dots, \partial_m)$. By continuity, this remains true on some neighborhood of p . Therefore

we have a non-singular projection $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^k$ onto the coordinate hyperplane of the coordinates (x^1, \dots, x^k) . Then $\pi_*|_D$ is a smooth bundle isomorphism.

Thus the vector fields $\pi^*(\partial_i)$, $i = 1, \dots, k$, locally span D . By properties of Lie derivatives, $[\pi^*(\partial_i), \pi^*(\partial_j)] = \pi^*[\partial_i, \partial_j] = 0$. Thus D is indeed spanned by a holonomic k -frame and therefore integrable. ■

This theorem is of crucial importance in the theory of Lie groups, fiber bundles, and even dynamical systems. The entire concept of *integrability* ultimately stems from it. The rest of this section is devoted to characterizing integral submanifolds and proving a global version of the Frobenius theorem, which will be useful in some applications.

Corollary 9.5.5 ([7, Cor. 19.13]). *Let D be a rank k involutive distribution on M . If $S \lhd M$ is a codimension k embedded submanifold such that $T_p S \cap D_p = \{0\}$ for all $p \in S$, then there is a flat chart $(U, \{s^i\})$ for S centered at p in which $S \cap U$ is the slice $s^1 = \dots = s^k = 0$.*

Proof. The proof of Frobenius theorem showed that locally D is spanned by k commuting vector fields, so this Corollary follows by the Rectification Theorem 9.4.8. ■

Proposition 9.5.6 ([7, Prop 19.16]). *Let D be an involutive distribution of rank k on a smooth manifold M , and let $(U, \{s^i\})$ be a flat chart for D . If H is an integral manifold of D , then $H \cap U$ is a union of countably many disjoint open subsets of parallel k -dimensional slices of U , each of which is open in H and embedded in M .*

Proof. Since the inclusion $i : H \hookrightarrow M$ is continuous, $H \cap U = i^{-1}(U)$ is open in H and this is a countable disjoint union of connected components, each open in H .

Let V be a component of $H \cap U$. Since the chart is flat, the functions x^{k+1}, \dots, x^n are constant along D and therefore constant on any connected slice like V , and V lies in a single slice S .

Since S is embedded in M , the inclusion $V \hookrightarrow M$ also restricts to a smooth map into S . The inclusion $V \hookrightarrow S$ is thus an injective smooth immersion between manifolds of the same dimension, and therefore a local diffeomorphism, an open map, and a homeomorphism onto an open subset of S . The inclusion map $V \hookrightarrow M$ is a composition of the smooth embeddings $V \hookrightarrow S \hookrightarrow M$ and thus a smooth embedding. ■

Recall now that a weakly embedded submanifold H is such that any map into M whose values fall into H restricts to a smooth map into H .

Theorem 9.5.7 ([7, Thm. 19.17] or [16, Prop. 3.5.15]). *Every integral manifold of an involutive distribution is weakly embedded.*

Proof. Let $H \subset M$ be an integral manifold of an involutive rank- k distribution D on M and let $F : N \rightarrow M$ be a smooth map such that $F(N) \subset H$. Choose $p \in H$ and set $q = F(p) \in H$. Let $\{y^i\}_{i=1}^n$ be flat coordinates for D on a neighborhood U of q , and let $\{x^i\}$ be smooth coordinates for N on a connected neighborhood B of p such that $F(B) \subset U$. With the coordinate representation of F written as

$$y^i = F^i(x), \quad (9.5.1)$$

the fact that $F(B) \subset H \cap U$ means that the functions F^{k+1}, \dots, F^n take on only countably many values. Since B is connected, these coordinate functions are constant, and thus $F(B)$ lies in a single slice $S \subset U$. Because $S \cap H$ is an open subset in H that is embedded in M , it follows that $F|_B$ is smooth from B into $S \cap H$, and thus by composition, $F|_B : B \rightarrow S \cap H \hookrightarrow H$ is smooth into H . ■

Corollary 9.5.8. *Images of the maximal integral curves of a vector field are weakly embedded submanifolds.*

Remark 9.5.9. The version of the Frobenius theorem proved above is local, i.e. integral submanifolds are guaranteed to exist only in small neighborhoods. However, the full version of the Frobenius theorem is much more powerful as it establishes a bijection between smooth integrable distributions and smooth *foliations* of M by maximal integral submanifolds (analogs of maximal integral curves of vector fields). A foliation is a decomposition of M into a disjoint union of immersed submanifolds, called *leaves*, such that every point has a local chart in which the foliation is flat (i.e. the leaves are just k -dimensional hyperplanes of the form $x^{k+1} = c^{k+1}, \dots, x^n = c^n$). For a global version of the Frobenius theorem, see [7, Ch. 19] or [16, § 3.5].

Lemma 9.5.10 ([16, Lem. 3.5.14]). *Let D be a distribution on M .*

1. *Let $N_1, N_2 \subset M$ be two integral manifolds of D . If $N_1 \cap N_2 \neq \emptyset$ then $N_1 \cap N_2$ is open in N_1 and N_2 and the smooth structures on it induced from either coincide. There is a unique smooth structure on $N_1 \cup N_2$ such that N_1 and N_2 are open submanifolds. W.r.t. this structure $N_1 \cup N_2$ is an integral manifold of D .*
2. *Let $N \subset M$ be an integral manifold of D , let $m \in M$ such that $\dim D_m = \dim N$ and let (U, κ) be a local flat chart for D at m . If a slice U_c of U intersects N , then U_c is an integral manifold of D and $N \cap U_c$ is open and closed w.r.t. the topology induced from N . In particular, the number of slices of U intersecting N is at most countable.*

Proof. 1. Let $\dim N_1 = \dim N_2 = r$. Choose a local r -frame $\{X_i\}_{i=1}^r$ for D at m . Since these vector fields are tangent to N_i , they are also local vector fields on N_i . Then the flows of these vector fields together generate a local coordinate system on N_i ,

- mapping $(-\epsilon, \epsilon)^r$ into an open subset of N_i . By construction these coordinates will coincide and therefore their values will cover the same subset $V \subset N_1 \cap N_2$ open in both N_i 's. W.r.t. the differentiable structure induced from either N_i , V is diffeomorphic to $(-\epsilon, \epsilon)^r$. Since the construction works for any $m \in N_1 \cap N_2$, the assertion follows.
2. Let $r = \dim N = \dim D_m$. If N intersects a slice U_c , then U_c contains a point where $\text{rank } D = r$. The U_c is an integral manifold of D by dimension-counting. Now assertion 1 implies that $N \cap U_c$ is open in N and hence in $N \cap U$. Since this holds for all slices intersecting N and since

$$N \cap U_c = (N \cap U) \setminus \left(\bigcup_{c' \neq c} N \cap U_{c'} \right), \quad (9.5.2)$$

the intersection $N \cap U_c$ is also closed in $N \cap U$. It follows that $N \cap U_c$ is a union of connected components of $N \cap U$. Since the slices are disjoint, this implies that the number of slices which intersect N is at most as large as the number of connected components of $N \cap U$. Since the topology on $N \cap U$ is induced from the manifold N , it is second countable, hence the number of connected components is at most countable. ■

Theorem 9.5.11 ([16, Thm. 3.5.17]). *Let D be an integrable distribution on M . For every $m \in M$ there exists a unique integral manifold N_m of D through m which is maximal in the following sense. For every integral manifold N of D such that $N \cap N_m \neq \emptyset$ it holds that $N \subset N_m$, and N is an open submanifold of N_m .*

Proof. Let $r = \dim D_m$ and define a subset $N_m \subset M$ by taking the union of all integral manifolds of D through m . Equip it with the topology generated by the open subsets of these integral manifolds. The union of all maximal atlases of these manifolds also defines an atlas on N_m . By Lemma 9.5.10 this atlas is smooth. One can then show that N_m is second countable (see [7, Lem. 19.22] or [35, Thm. 1.64]). Then N_m is a manifold of dimension r . By construction, the local representatives of the inclusion $N_m \hookrightarrow M$ are smooth, hence N_m is a smooth submanifold of M . Also by construction, every $\tilde{m} \in N_m$ belongs to an integral manifold N through m and N is an open submanifold of N_m , so N_m is also an integral manifold through m . It has the maximality property stated in the theorem because if some integral manifold N of D intersects N_m , then $N \cup N_m$ is also integral and thus already contained in N_m . It follows that $N \cap N_m = N$ and hence N is an open submanifold of N_m . Uniqueness follows because any two maximal integral manifolds through m would coincide as sets by definition and each of them would be an open submanifold of the other, hence they would also coincide as manifolds. ■

Definition 9.5.12 (Foliation). A foliation of an n -dimensional smooth manifold M is a family \mathcal{N} of connected submanifolds of M , called the *leaves* of the foliation, such that

1. the leaves are disjoint and $\cup_{N \in \mathcal{N}} N = M$,
2. for every $m \in M$ there exists a local flat chart (U, κ) of M satisfying
 - (a) $\kappa(m) = 0$ and $\kappa(U) = (-\epsilon, \epsilon)^n$ for some $\epsilon > 0$,
 - (b) the leaves are invariant under the flows of the local coordinate vector fields $\partial_1, \dots, \partial_r$, where r denotes the dimension of the leaf containing m , also called the rank of \mathcal{N} at m .

A foliation is called *regular* if its rank is constant, and singular otherwise.

Theorem 9.5.13 (Global Frobenius Theorem [16, Prop. 3.5.21] or [7, Thm. 19.21]). *The collection of all maximal connected integral manifolds of an involutive distribution D on a smooth manifold M forms a foliation of M .*

The assignment of the family of maximal integral manifolds to an integrable distribution defines a bijection between smooth integrable distributions on M and smooth foliations of M . Regular distributions of rank r thereby correspond to regular foliations of rank r .

Proof. Theorem 9.5.11 implies that condition 1 of the definition of a foliation is satisfied. The invariance of maximal integral manifolds under the flows of local D -valued vector fields lets us generate local flat flat charts via a composition (as in rectification for local frames) of the flows and hence condition 2 is satisfied as well.

Conversely, let \mathcal{N} be a smooth foliation. For $m \in M$, let D_m be the tangent space at m of the leaf containing m . This defines a subset $D \subset TM$. By definition of a foliation, D_m is spanned by the values of local D -valued vector fields at m . Hence D is a distribution. By construction, every leaf is an integral manifold of D . First, this implies that D is integrable. Second, in view of Theorem 9.5.11, this implies that every leaf of \mathcal{N} is contained in an open subset in a maximal integral manifold of D . Thus \mathcal{N} is a union of leaves. Since the leaves are disjoint and \mathcal{N} is connected, \mathcal{N} must coincide with a single leaf. This shows that \mathcal{N} is the family of maximal integral manifolds of D , and the assignment is bijective, indeed. ■

Remark 9.5.14. Overdetermined systems of linear PDEs are solvable iff the distribution spanned by the differential operators as vector fields is integrable:

$$\begin{cases} X_1 \cdot u(x) = f_1(x), \\ \dots \\ X_k \cdot u(x) = f_k(x) \end{cases} \quad \text{has solutions iff } \exists C_{ij}^k \in C^\infty \text{ s.t. } \begin{cases} [X_i, X_j] = C_{ij}^k X_k, \\ X_i \cdot f_j - X_j \cdot f_i = C_{ij}^k f_k. \end{cases} \quad (9.5.3)$$

Example 9.5.15. Show that the distribution on $\mathbb{R}^3 \setminus \{0\}$ generated by the vector fields

$$X_1 = x_2 \partial_3 - x_3 \partial_2, \quad X_2 = x_3 \partial_1 - x_1 \partial_3, \quad X_3 = x_1 \partial_2 - x_2 \partial_1, \quad (9.5.4)$$

is regular of rank 2 and involutive. Find the maximal integral manifolds. How are they related to the action of $\text{SO}(3)$ on \mathbb{R}^3 ?

10 Differential Forms and Integration

10.1 Exterior algebra

Definition 10.1.1 (Exterior forms). Let V be a \mathbb{K} -vector space. The alternation and symmetrization maps are defined on the tensor space $T^k V = (V)^{\otimes k}$ respectively as

$$\text{Alt } t = \frac{1}{k!} \sum_{\sigma \in S_k} (\text{sign } \sigma) \cdot t^\sigma, \quad (10.1.1)$$

$$\text{Sym } t = \frac{1}{k!} \sum_{\sigma \in S_k} t^\sigma, \quad (10.1.2)$$

where the operation t^σ is defined on a basis as $(e_{i_1} \otimes \cdots \otimes e_{i_k})^\sigma = e_{\sigma(i_1)} \otimes \cdots \otimes e_{\sigma(i_k)}$ and extended to all of $T^k V$ by linearity. We also denote $\wedge^k V = \text{Alt}(T^k V)$ and $\Sigma^k V = \text{Sym}(T^k V)$. Elements of the space $\wedge^k V^*$ are called *exterior forms* of degree k on V . If $\dim V = n$, then the dimension of this space is $\dim \wedge^k V = \binom{n}{k}$. In particular, $\wedge^1 V = V$, $\dim \wedge^0 V = \dim \wedge^n V = 1$ and $\dim \wedge^{>n} V = 0$.

Both \wedge^k and Σ^k are covariant functors $\text{FinVect}_{\mathbb{K}} \rightarrow \text{FinVect}_{\mathbb{K}}$ that act on morphisms by restrictions $\wedge^k f = f^{\otimes k}|_{\wedge^k V}$ and $\Sigma^k f = f^{\otimes k}|_{\Sigma^k V}$.

If V has Cartesian coordinates x^1, \dots, x^n , then a basis of V is formed by $\{\partial_{x^i}\}$. The dual basis of V^* is conventionally denoted by $\{dx^i\}$, so that

$$dx^i(\partial_{x^j}) = \delta_j^i. \quad (10.1.3)$$

Then this basis induces a basis of $\wedge^k V^*$ consisting of the $\binom{n}{k}$ elements

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} := k! \text{Alt}(dx^{i_1} \otimes \cdots \otimes dx^{i_k}), \quad 1 \leq i_1 < i_2 < \cdots < i_k \leq n. \quad (10.1.4)$$

Together with the wedge product, the direct sum of all exterior powers of V makes up an algebra, called the exterior, or Grassmann, algebra of V :

$$\bigwedge V = \bigoplus_{k=0}^{\infty} \bigwedge^k V. \quad (10.1.5)$$

Definition 10.1.2 (Determinant of a linear map). Let V be a vector space over a field \mathbb{K} with $\dim V = n$ and let $f \in \text{End}(V)$. The determinant of f is defined as

$$\det f := \bigwedge^n f. \quad (10.1.6)$$

Being an element of $\text{Hom}(\bigwedge^n V, \bigwedge^n V)$, which is naturally isomorphic to the field \mathbb{K} , it is a well-defined scalar.

Proposition 10.1.3. 1. If $\dim V = n$, then for any $t \in \bigwedge^n V^*$ and any $f \in \text{End}(V)$,

$$f^* t = \det f \cdot t. \quad (10.1.7)$$

2. If $f \in \text{End}(V)$ is represented by a matrix $f(e_i) = A_i^j e_j$ in some basis $\{e_i\}$ of V , then

$$\det f = \det A = \sum_{\sigma \in S_n} (\text{sign } \sigma) A_{\sigma(1)}^1 \cdots A_{\sigma(n)}^n. \quad (10.1.8)$$

Proof. 1. By definition of f^* , if $v \in \bigwedge^n V$, then

$$(f^* t)(v) = t(f(v)) = t(\det f \cdot v) = \det f \cdot t(v). \quad (10.1.9)$$

2. A basis element of $\bigwedge^n V$ is $e_1 \wedge \cdots \wedge e_n$. Acting on it, $\det f = \bigwedge^n f$ gives

$$\begin{aligned} & \left(\bigwedge^n f \right) (e_1 \wedge \cdots \wedge e_n) = k! f^{\otimes n} \circ \text{Alt}(e_1 \otimes \cdots \otimes e_n) = k! \text{Alt}(f(e_1) \otimes \cdots \otimes f(e_n)) = \\ &= k! \text{Alt}(A_1^{j_1} \cdots A_n^{j_n} \cdot e_{j_1} \otimes \cdots \otimes e_{j_n}) = \sum_{\pi \in S_n} (\text{sign } \pi) \cdot A_{\pi(1)}^{j_1} \cdots A_{\pi(n)}^{j_n} \cdot e_{j_1} \otimes \cdots \otimes e_{j_n} = \\ &= \sum_{\sigma, \pi \in S_n} (\text{sign } \sigma \cdot \text{sign } \pi) \cdot A_{\pi(1)}^1 \cdots A_{\pi(n)}^n \cdot e_{\sigma(1)} \otimes \cdots \otimes e_{\sigma(n)} = \\ &= \left(\sum_{\pi \in S_n} (\text{sign } \pi) A_{\pi(1)}^1 \cdots A_{\pi(n)}^n \right) \cdot e_1 \wedge \cdots \wedge e_n. \end{aligned} \quad (10.1.10)$$

By item 1, the expression in the brackets is the determinant. ■

Definition 10.1.4 (Exterior/wedge product). For $\alpha \in T_p^0 V$ and $\beta \in T_q^0 V$, we define their exterior/wedge product by

$$\alpha \wedge \beta = \frac{(p+q)!}{p!q!} \text{Alt}(\alpha \otimes \beta) \in \bigwedge^{p+q} V. \quad (10.1.11)$$

Wedge products have the following basic properties:

1. $\alpha \wedge \beta = \text{Alt}(\alpha) \wedge \beta = \alpha \wedge \text{Alt}(\beta)$;
2. \wedge is bilinear and $\alpha \wedge \beta = (-1)^{\deg \alpha \cdot \deg \beta} \beta \wedge \alpha$;

$$3. \alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma.$$

4. For $\alpha_i \in V^*, i = 1, \dots, k$,

$$\begin{aligned} (\alpha_1 \wedge \cdots \wedge \alpha_k)(v_1, \dots, v_k) &= \sum_{\sigma \in S_k} (\text{sign } \sigma) \cdot \alpha_1(v_{\sigma(1)}) \cdots \alpha_k(v_{\sigma(k)}) = \\ &= \det \begin{bmatrix} \alpha_1(v_1) & \cdots & \alpha_1(v_k) \\ \vdots & \ddots & \vdots \\ \alpha_k(v_1) & \cdots & \alpha_k(v_k) \end{bmatrix}. \quad (10.1.12) \end{aligned}$$

5. For $\alpha, \beta \in \bigwedge^\bullet V^*$ and any linear map $f : V' \rightarrow V$,

$$f^*(\alpha \wedge \beta) = (f^*\alpha) \wedge (f^*\beta). \quad (10.1.13)$$

Definition 10.1.5 (Exterior powers of VBs). Following the logic of Definition 8.1.10, we can extend the functors \bigwedge^p and Σ^p to all smooth vector bundles. We focus only on \bigwedge^p . For a vector bundle $E \xrightarrow{\pi} M$, we obtain the exterior powers

$$\bigwedge^p E = (E)_{\text{asym}}^{\otimes p}. \quad (10.1.14)$$

p is called the *degree* of the elements or sections of $\bigwedge^p E$:

$$\deg \alpha = p \text{ for } \alpha \in \Gamma \left(\bigwedge^p E \right). \quad (10.1.15)$$

The wedge product \wedge gets extended to sections of these bundles fiberwise. In particular, if $\text{rank } E = k$, then

$$\det E := \bigwedge^k E \quad (10.1.16)$$

is called the *determinant (line) bundle* of E . If the transition functions of E are $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow GL_k(\mathbb{R})$, then the transition functions of $\det E$ are $\det \tau_{\alpha\beta}$.

The total rank of the exterior bundles is $\text{rank} \left(\bigoplus_{p \geq 0} \bigwedge^p E \right) = 2^k$.

The alternation and symmetrization maps get extended to $\text{Alt} : T^p E \rightarrow \bigwedge^p E$ and $\text{Sym} : T^p E \rightarrow \Sigma^p E$.

All properties of wedge products listed above immediately generalize to sections of the exterior bundles.

Definition 10.1.6 (Determinant of a bundle map). If $E \xrightarrow{\pi} M$ is a **VB** of rank k and $f : E \rightarrow E$ is a bundle map, then we define $\det f = \bigwedge^k f$. If $\omega \in \Gamma(\det E^*)$, then $f^*\omega = \det f \cdot \omega$.

Definition 10.1.7 (Exterior/Grassmann algebra). For a **VB** E , the space $\bigoplus_{p \geq 0} \Gamma(\bigwedge^p E) = \Gamma\left(\bigoplus_{p \geq 0} \bigwedge^p E\right)$ with the product \wedge is called the exterior, or Grassmann, algebra of E .

Exercise 10.1. 1. Prove formula 10.1.12.

2. Show that k one-forms (or vectors) $\{\alpha_i\}$ are linearly independent iff $\alpha^1 \wedge \cdots \wedge \alpha^k \neq 0$.

10.2 Differential forms

Definition 10.2.1 (Differential forms). Differential forms of degree p on a smooth manifold M are sections of $\bigwedge^p(T^*M)$. The exterior algebra of differential forms is denoted

$$\Omega^p(M) := \Gamma\left(\bigwedge^p(T^*M)\right). \quad (10.2.1)$$

In particular, we have $\Omega^0(M) = C^\infty(M)$, $\Omega^1(M) = \mathfrak{X}^*(M)$, and $\Omega^p(M) = \bigwedge^p \mathfrak{X}^*(M)$.

Recall that for $f \in C^\infty(M)$, $df = \text{pr}_2 \circ f_*$, where pr_2 is the projection onto the fiber in $\mathbb{TR} = \mathbb{R} \times \mathbb{R}$. As we have established in Example 8.3.11 using the tensor characterization theorem, the differential is a one-form, $df \in \Omega^1(M)$.

Proposition 10.2.2. If $x^i \in C^\infty(U)$, $i = 1, \dots, m$ are local coordinate functions on $U \subset M^m$, then the differentials dx^i of these functions coincide with the dual coordinate frame, i.e.

$$dx^i(\partial_j) = \delta_j^i. \quad (10.2.2)$$

Proof. By definition, of components of vector fields as push-forwards by x_* ,

$$dx^i(\partial_j) = \text{pr}_2 \circ x_*^i(\partial_j) = (\partial_j)^i = \delta_j^i. \quad (10.2.3)$$

Corollary 10.2.3. Locally, the differential df of a function $f \in C^\infty(M)$ is the standard “total differential” from multivariable calculus,

$$df = df(\partial_i) \cdot dx^i = (\partial_i f) \cdot dx^i. \quad (10.2.4)$$

Exercise 10.2. On \mathbb{R}^2 with Cartesian coordinates (x, y) , compute $(dx \wedge dy)(\nabla f, \nabla g)$ for any $f, g \in C^\infty(\mathbb{R}^2)$. Compare to $(df \wedge dg)(\partial_x, \partial_y)$ and $(\partial_x \wedge \partial_y)(df, dg)$ (where we act with

vectors on covectors via the canonical pairing between TM and T^*M).

Theorem 10.2.4. *There exists a unique family of differential operators $\tilde{d}_p(U) : \Omega^p(U) \rightarrow \Omega^{p+1}(U)$ for every $U \subset M$, denoted collectively by \tilde{d} , such that:*

1. \tilde{d} is an \mathbb{R} -linear \wedge -derivation: $\tilde{d}(\alpha \wedge \beta) = \tilde{d}\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge \tilde{d}\beta$;
2. $\forall f \in C^\infty(M), \tilde{d}f = df$;
3. $\tilde{d} \circ \tilde{d} = 0$;
4. \tilde{d} is local: $\tilde{d}(\alpha|_U) = \tilde{d}\alpha|_U$.

Proof. Suppose the existence of \tilde{d} is established. Then let us consider its action on differential forms of the form $f_0 \cdot df_1 \wedge \cdots \wedge df_p$ where $f_i \in C^\infty(M)$ (such forms span the entire space $\Omega^p(M)$). Using the properties (1)-(3), we find

$$\tilde{d}(f_0 \cdot df_1 \wedge \cdots \wedge df_p) \stackrel{(1),(3)}{=} (\tilde{d}f_0) \wedge df_1 \wedge \cdots \wedge df_p \stackrel{(2)}{=} (\partial_{i_0} f_0) \cdot dx^{i_0} \wedge df_1 \wedge \cdots \wedge df_p. \quad (10.2.5)$$

This formula in fact completely determines the action of \tilde{d} on all differential forms! Namely, if locally $\omega = \sum_{1 \leq i_1 < \dots < i_p \leq m} \omega_{i_1 i_2 \dots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ with $\omega_{i_1 \dots i_p} \in C^\infty(U)$, then

$$\tilde{d}\omega = \sum_{i_0=1}^m \sum_{1 \leq i_1 < \dots < i_p \leq m} \partial_{i_0} \omega_{i_1 i_2 \dots i_p} \cdot dx^{i_0} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}. \quad (10.2.6)$$

This formula is in fact coordinate-independent and therefore defines a true differential operator on the algebra of differential forms. All four properties immediately follow (Exercise; only property (3) is nontrivial). This proves the existence and uniqueness. ■

Definition 10.2.5 (Exterior derivative). The differential operator whose existence and uniqueness is established in Theorem 10.2.4 is called the exterior derivative of differential forms and is denoted just $d : \Omega^p(M) \rightarrow \Omega^{p+1}(M)$.

Proposition 10.2.6. *If $F \in C^\infty(M, N)$, then $F^* : \Omega(N) \rightarrow \Omega(M)$ is a homomorphism of algebras and*

$$F^* d\omega = dF^* \omega. \quad (10.2.7)$$

Sometimes this property is described as “differentials are invariant under changes of variables”.

Proof. The homomorphism part follows from (10.1.13). The second part can be checked in local coordinates using the homomorphism property. First, for $f \in C^\infty(M)$, using Proposition

9.3.13 (note that F doesn't really have to be a diffeomorphism when the pullback is computed at a separate point),

$$(F^*df)(X)(p) = df(F_*X_p) = (F^*\mathcal{L}_{F_*X}f)(p) = (\mathcal{L}_X F^*f)(p) = (dF^*f)(p), \quad (10.2.8)$$

next we can apply this to every component of a differential form in local coordinates, since these components are functions:

$$d(F^*\omega) = F^*(d\omega_{i_1 \dots i_p} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_p}) = (F^*d\omega_{i_1 \dots i_p}) \wedge F^*dx^{i_1} \wedge \dots \wedge F^*dx^{i_p} = F^*(d\omega). \quad (10.2.9)$$

■

Corollary 10.2.7. $\Omega : \text{Man}^\infty \rightarrow \text{Alg}$ is a functor, and in fact it is a functor into the category of commutative differential graded algebras.³⁰

Definition 10.2.8 (Inner product). Let $\omega \in \Omega^{p+1}(M)$, $p \geq 0$, and $X \in \mathfrak{X}(M)$. The inner product with X is the operator $i_X : \Omega^{p+1}(M) \rightarrow \Omega^p(M)$ defined by

$$(i_X\omega)(X_1, \dots, X_p) = \omega(X, X_1, \dots, X_p). \quad (10.2.10)$$

We also put $i_X|_{\Omega^0(M)} = 0$.

Proposition 10.2.9. For any $X \in \mathfrak{X}(M)$ and any $\omega \in \Omega^\bullet(M)$,

1. $d\mathcal{L}_X = \mathcal{L}_X d$ on $\Omega^\bullet(M)$;
2. i_X is a \wedge -derivation on the exterior algebra;
3. $i_{f \cdot X} = f \cdot i_X$;
4. $i_X df = \mathcal{L}_X f$ for $f \in \Omega^0(M)$;
5. $\mathcal{L}_X|_{\Omega^p(M)} = i_X \circ d + d \circ i_X$ (Cartan's magic formula);
6. $\mathcal{L}_{fX}\omega = f\mathcal{L}_X\omega + df \wedge i_X\omega$;
7. For any diffeo F , $F^*i_X\omega = i_{F^*X}F^*\omega$.

Proof. 1. First we check that the Lie derivative of a differential form is a differential form. It suffices to check the property on differential forms of the form $\alpha^1 \wedge \dots \wedge \alpha^p$, $\alpha^i \in \Omega^1(M)$. We know that \mathcal{L}_X is a tensor derivation, so

$$\mathcal{L}_X(\alpha^1 \wedge \dots \wedge \alpha^p) = \sum_{i=1}^p \alpha^1 \wedge \dots \wedge \alpha^{i-1} \wedge (\mathcal{L}_X \alpha^i) \wedge \alpha^{i+1} \wedge \dots \wedge \alpha^p. \quad (10.2.11)$$

Since $\Omega^1(M) = \mathfrak{X}(M)$, we know that $\mathcal{L}_X \alpha^i$ are also one-forms, and thus \mathcal{L}_X is a differential form.

Furthermore, if F denotes the flow of X , then $\mathcal{L}_X\alpha(p) = \frac{d}{d\lambda}|_{\lambda=0} F^*\alpha$. But we already know that d commutes with pullbacks F_λ^* , so the result follows;

2. Exercise (direct check by applying the definitions to a basis);
3. Follows from the linearity of i_X ;
4. This is literally the definition of \mathcal{L}_X ;
5. Proof by induction in $p = \deg \omega$. The case $p = 0$ reduces to part (4). Any $(p+1)$ -form can be locally represented as $\sum_i df_i \wedge \omega_i$ with $f_i \in \Omega^0(U)$, $\omega_i \in \Omega^p(U)$. Thus we need to check the property only for forms $df \wedge \omega$:

$$\mathcal{L}_X(df \wedge \omega) = \mathcal{L}_X df \wedge \omega + df \wedge \mathcal{L}_X \omega \quad (10.2.12)$$

and

$$\begin{aligned} i_X d(df \wedge \omega) + d i_X(df \wedge \omega) &= -i_X(df \wedge d\omega) + d(i_X df \wedge \omega - df \wedge i_X \omega) \stackrel{(2)}{=} \\ &= \cancel{-i_X df \wedge d\omega} + df \wedge i_X d\omega + d \underbrace{i_X df \wedge \omega}_{\mathcal{L}_X f} + \cancel{i_X df \wedge d\omega} + df \wedge di_X \omega \stackrel{\text{induction}}{=} \\ &= df \wedge \mathcal{L}_X \omega + d \mathcal{L}_X f \wedge \omega \stackrel{(4)}{=} \mathcal{L}_X(df \wedge \omega). \quad (10.2.13) \end{aligned}$$

6. Follows immediately from part 5.
7. Explicit check using the definitions of i_X and F^* by acting with both sides on a set of vector fields.

■

Willmore's theorem implies

$$(\mathcal{L}_X \omega)(X_1, \dots, X_p) = \mathcal{L}_X(\omega(X_1, \dots, X_p)) - \sum_{i=1}^p \omega(X_1, \dots, [X, X_i], \dots, X_p). \quad (10.2.14)$$

We use this property in the proof of the following useful formula.

Proposition 10.2.10 ([16, Prop. 4.1.6]). *The following explicit coordinate-independent formula for the exterior derivative holds:*

$$\begin{aligned} (d\omega)(X_0, \dots, X_p) &= \sum_{i=0}^p \mathcal{L}_{X_i}(\omega(X_0, \dots, \hat{X}_i, \dots, X_p)) + \\ &+ \sum_{0 \leq i < j \leq p} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_p), \quad (10.2.15) \end{aligned}$$

where hats indicate that the corresponding vector fields are excluded from the list of arguments. In particular, for $\alpha \in \Omega^1(M)$,

$$(\mathrm{d}\alpha)(X_1, X_2) = X_1 \cdot \alpha(X_2) - X_2 \cdot \alpha(X_1) + \alpha([X_1, X_2]). \quad (10.2.16)$$

Proof. Proof by induction in p . For $p = 0$, we have already established this formula, $\mathrm{d}\omega(X_0) = \mathcal{L}_{X_0}\omega$.

For $p > 0$, $(\mathrm{d}\omega)(X_0, \dots) = (i_{X_0}\mathrm{d}\omega)(X_1, \dots) = (\mathcal{L}_{X_0}\omega - \mathrm{d}i_{X_0}\omega)(X_1, \dots)$, and by the induction step at $p - 1$, we can expand $\mathrm{d}i_{X_0}\omega$ and permute the result to bring it to the required form. ■

Exercise 10.3. 1. Show that if $f_i \in C^\infty(M)$, $i = 1, \dots, k$ are *functionally dependent*, i.e. there exists a $g \in C^\infty(\mathbb{R}^n)$ such that $g(f_1, \dots, f_k) = 0$, then $\mathrm{d}f_1 \wedge \dots \wedge \mathrm{d}f_k = 0$.
2. What conditions need to be added to make part 1 hold in the reverse direction?

10.3 Orientation

The group $\mathrm{GL}_n^+(\mathbb{R})$ is defined as the connected component of the identity in $\mathrm{GL}_n(\mathbb{R})$, i.e. the group of matrices with a strictly positive determinant.

Definition 10.3.1 (Orientation on a vector space). Let V be an n -dimensional real vector space. Two bases $\{e_i\}$ and $\{e'_j\}$ are said to *fix the same orientation* on V if there exists a matrix $A \in \mathrm{GL}_n^+(\mathbb{R})$ such that $e_i = \sum_{j=1}^n A_i^j e'_j$, $1 \leq i \leq n$. The two equivalence classes of like-oriented bases on V form the factor-set $O(V)$. An orientation on V is a choice of an element $\mu \in O(V)$. Any basis belonging to the chosen orientation, $\{e_i\} \in \mu$, is called *positively oriented* w.r.t. that orientation.

Note that even though the set of possible orientations has two elements, there is no way to naturally identify it with the group \mathbb{Z}_2 , because neither of the orientations is distinguished.

Since $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n^+(\mathbb{R}) = \mathbb{Z}_2$, we have a natural group homomorphism $\vartheta : \mathrm{GL}_n(\mathbb{R}) \rightarrow \mathbb{Z}_2 = \{-1, +1\}$ given explicitly by $\vartheta(A) = \mathrm{sign}(\det A)$. The action of the group $\mathrm{GL}(V)$ on the set of bases on V then induces a natural action of the group \mathbb{Z}_2 on the set $O(V)$. Namely, if $A \in \mathrm{GL}(V)$ and $\{e_i\}$ is a basis, then $\{A \cdot e_i\}$ is oriented like $\{e_i\}$ if $\vartheta(A) = 1$, and it has the opposite orientation if $\vartheta(A) = -1$. Therefore, the generator of the group \mathbb{Z}_2 simply swaps the two elements of $O(V)$.

Note that any *linear isomorphism* $f : V \rightarrow V'$ induces a map $O(f) : O(V) \rightarrow O(V')$, namely $O(f)(\mu) = \mu'$ if for any basis $\{e_i\} \in \mu$ the basis $\{f(e_i)\}$ belongs to μ' . Therefore we have constructed a covariant functor $O : \mathrm{FinVect}_{\mathbb{R}} \rightarrow \mathrm{2Set}$, where $\mathrm{2Set}$ is the category of two-element sets.

Definition 10.3.2 (Orientation-preserving linear maps). Let (V, μ) and (V, μ') be two oriented vector spaces. A linear map $f \in \text{Hom}(V, V')$ is called orientation-preserving if it maps any positively oriented basis on V to a positively oriented basis on V' , i.e. it is an isomorphism and $O(f)(\mu) = \mu'$.

Definition 10.3.3 (Orientation bundle). Let $E \xrightarrow{\pi} M$ be a real vector bundle with the typical fiber F . The orientation bundle of E is the double cover $O(E) \xrightarrow{\pi} M$ (i.e. a **FB** whose typical fiber consists of two points) obtained by the application of the smooth functor O fiberwise as in Definition 8.1.10. Namely, if $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \text{GL}(F)$ is a cocycle of E , then a cocycle of $O(E)$ is $\vartheta \circ \tau_{\alpha\beta} = \text{sign det } \tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \mathbb{Z}_2$.

Exercise 10.4. Show that the orientation bundle of the Möbius band is isomorphic (as a fiber bundle) to the connected double cover of the circle.

Definition 10.3.4 (Orientation on a bundle, orientation-preserving maps). A real vector bundle $E \xrightarrow{\pi} M$ is called orientable if its orientation bundle $O(E)$ is trivial. An orientation on such a bundle is a choice of one of the two connected components of $O(E)$. A manifold M is called orientable if its tangent bundle TM is orientable. An orientation on M is an orientation on TM .

A bundle map $f : E \rightarrow E'$ between two oriented bundles is called *orientation-preserving* if it preserves the orientation of every fiber. A smooth map $f : M \rightarrow N$ between manifolds is orientation-preserving if $f_* : TM \rightarrow TN$ is.

Proposition 10.3.5. TFAE:

1. A real **VB** $E \xrightarrow{\pi} M$ is orientable;
2. There exists a global section of the orientation bundle $O(E)$;
3. The determinant line bundle $\det E$ is trivial;
4. There exists a non-vanishing section of the determinant line bundle $\det E$;
5. There exists a bundle atlas on E such that the corresponding cocycle $\{\tau_{\alpha\beta}\}$ has $\det \tau_{\alpha\beta} > 0$, i.e. if the structure group of E can be reduced to $\text{GL}_k^+(\mathbb{R})$.

Proof. Exercise. ■

Exercise 10.5. 1. Show that the Möbius band is non-orientable as a line bundle over the circle.

2. Show that the Möbius band is non-orientable as a manifold.

Exercise 10.6. 1. Show that the tangent bundle TM of *any* manifold is orientable as a manifold.

2. Show that any parallelizable manifold is orientable.

3. Show that an orientable **VB** over an orientable base is orientable as a manifold. Give an

example of the converse not being true.

The following notions can be defined for general vector bundles too, but we will focus on tangent bundles.

Definition 10.3.6 (Volume form). A volume form on a manifold M^m is a nowhere vanishing differential form of the top degree $\mu \in \Omega^m(M)$. Given an orientation on M , μ is called *positive* if for any positively oriented local frame $\{e_i\}$ one has $\mu(e_1, \dots, e_n) > 0$.

Proposition 10.3.7.

1. A manifold M is orientable iff there exists a volume form on M ;
2. A choice of a volume form naturally induces a choice of an orientation;
3. Two volume forms μ, μ' on M fix the same orientation iff there exists $f \in C^\infty(M)$ such that $f > 0$ everywhere and $\mu = f \cdot \mu'$.

Proof. Exercise. ■

Definition 10.3.8. Let M and M' be oriented smooth manifolds with chosen volume forms μ, μ' , respectively. For a smooth map $f : M \rightarrow M'$, we define its determinant with respect to these volume forms by the equality

$$f^* \mu' = \left(\det_{\mu, \mu'} f \right) \cdot \mu. \quad (10.3.1)$$

We will usually suppress the dependence on the volume forms. Clearly f is orientation-preserving iff $\det f > 0$ everywhere, and f is a local diffeo iff $\det f \neq 0$ everywhere.

The determinant is still multiplicative: $\det(f \circ g)(p) = (\det f)(g(p)) \cdot \det g(p)$.

Definition 10.3.9 (Tensor densities). One can generalize the notion of a tensor field in the following way. Let $w \in \mathbb{Z}$ and consider the bundle $(\det M)^{\otimes w}$, where a negative power means $(\det M)^{* \otimes |w|}$. Then the sections of the bundle $T'_s M \otimes (\det M)^{\otimes w}$ (it has the same rank as $T'_s M$) are called tensor densities of weight w . An alternative definition uses the bundle $|\det M|^w$ whose transition functions are $|\det \tau_{\alpha\beta}|^w$ if $\tau_{\alpha\beta}$ are the transition functions of TM , and w in this case can be any real number $w \in \mathbb{R}$. In addition, one can introduce “pseudotensor densities” whose transition functions include $(\text{sign} \det \tau_{\alpha\beta}) \cdot |\det \tau_{\alpha\beta}|^w$.

Theorem 10.3.10. Let $S < M$ be an immersed submanifold of codim $S = 1$, let $N \in \Gamma(TM|_S)$ be a vector fields such that $N \nparallel S$ everywhere, and let M be oriented. Then there exists a unique orientation on S such that a frame (e_1, \dots, e_{n-1}) is positively oriented in TS iff (N, e_1, \dots, e_{n-1}) is oriented in $TM|_S$.

Proof. Let μ be a positive volume form on M . The it is easy to check that $\mu_S = i_N \mu$ is a volume form on S that defines an orientation as stated. ■

Corollary 10.3.11. Let M be orientable and let a function $f \in C^\infty(M)$ have a regular value $c \in \mathbb{R}$. Then the submanifold $S = f^{-1}(c) \lhd M$ is orientable.

Proof. Embed everything into \mathbb{R}^n and define the vector field $N = \nabla f$. Then $N \perp S$ and $N \neq 0$ on S because $f|_S$ is regular. Therefore N induces an orientation on S by the theorem above. ■

Corollary 10.3.12. The boundary ∂M of any orientable manifold with boundary is orientable. The standard orientation is fixed by outward normals.

Proof. Consider a collar neighborhood of ∂M and the outward vector field N on it. N thus induces an orientation on ∂M . This orientation is independent of the choice of N because its n -th component at the boundary is necessarily negative and is the only variable between different choices of N . ■

Definition 10.3.13 (Divergence of a vector field). Let μ be a volume form on M and $X \in \mathfrak{X}(M)$. The divergence of X with respect to μ is defined by the equality

$$\mathcal{L}_X \mu = (\operatorname{div}_\mu X) \cdot \mu. \quad (10.3.2)$$

X is called *incompressible/volume-preserving* if $\operatorname{div}_\mu X = 0$ (or if $(F_\lambda^X)^* \mu = \mu$ where F^X is the flow of X).

10.4 Integration

First we define integrals of top forms (i.e. forms of the top degree) on \mathbb{R}^m .

Definition 10.4.1. Let $U \subset \mathbb{R}^m$ and $\omega \in \Omega^m(U)$. This differential form can be written as $\omega = f \cdot dx^1 \wedge \cdots \wedge dx^m$ with $f \in C^\infty(U)$. We define its integral as

$$\int_U \omega := \int_U f dx^1 \wedge \cdots \wedge dx^m, \quad (10.4.1)$$

where the integral on the right-hand side is the standard Riemann integral of a smooth function. A similar definition is given to integrals of top forms on $U \subset \mathbb{R}_+^m$.

Now we can use local charts and the invariance of differentials to define integrals of top forms on oriented manifolds.

Definition 10.4.2. Let M^m be an oriented manifold (with or without a boundary) and let $\{U_\alpha, \varphi_\alpha\}$ be a positively oriented atlas for it (i.e. all φ_α 's are orientation-preserving, where

the standard orientation on \mathbb{R}^m is assumed). If $\omega \in \Omega^m(M)$ and $\text{supp } \omega \subset U_\alpha$, then we define

$$\int_M \omega = \int_{U_\alpha} \omega := \int_{\varphi_\alpha(U_\alpha)} (\varphi_\alpha)_* \omega. \quad (10.4.2)$$

Finally, letting $\{\chi_\alpha\}$ be a **POU** subordinate to the atlas, we define for all $\omega \in \Omega^m(M)$

$$\int_M \omega := \sum_\alpha \int_{U_\alpha} \chi_\alpha \cdot \omega. \quad (10.4.3)$$

Also, if the dimension is zero, $m = 0$, we define integrals of functions (zero-forms) on M as $\int_M f = \sum_{p \in M} (\pm f(p))$, where the sign in front of $f(p)$ depends on the orientation of the corresponding point.

Proposition 10.4.3. *The above definition is consistent, i.e. independent of the choice of atlas and of the choice of **POU**.*

Proof. The integral is independent of the atlas by invariance of integrals on \mathbb{R}^m . Namely, if ψ_α is an alternative chart on U_α , then

$$\int_{\varphi_\alpha(U_\alpha)} \varphi_{\alpha*} \omega = \int_{\psi_\alpha(U_\alpha)} \psi_{\alpha*} \omega \quad (10.4.4)$$

by the standard theorem from calculus: Riemann integrals are invariant under changes of variables (if the orientation is taken into account), in this case $x \mapsto \psi_\alpha \circ \varphi_\alpha^{-1}(x)$.

Furthermore, if ξ_α is a different **POU** subordinate to the same atlas, then

$$\sum_\alpha \int_{U_\alpha} \chi_\alpha \omega = \sum_\beta \sum_\alpha \int \xi_\beta \chi_\alpha \omega = \sum_\beta \int_{U_\beta} \xi_\beta \omega, \quad (10.4.5)$$

which proves independence on the **POU**.

Finally, it is easy to combine the two arguments above to show that the integral is invariant when both the atlas and the **POU** get replaced (Exercise). ■

Proposition 10.4.4 (Invariance of integrals). *For a diffeomorphism $f : M \rightarrow N$ between two oriented manifolds,*

$$\int_M f^* \omega = \text{sign}(\det f) \int_N \omega. \quad (10.4.6)$$

The sign of the determinant here is computed with respect to any two positive volume forms on M and N .

Proof. Exercise. ■

Theorem 10.4.5 (Stokes). *Let M^m be an oriented manifold with or without a boundary, let $\omega \in \Omega^{m-1}(M)$ be a form with compact support. Let $i : \partial M \hookrightarrow M$ be the embedding inclusion of the boundary. Then $i^*\omega \in \Omega^{m-1}(\partial M)$, which can be treated as the restriction of ω to ∂M , satisfies*

$$\int_{\partial M} i^*\omega = \int_M d\omega, \quad (10.4.7)$$

where the orientation on ∂M is induced from M by the outward normals.

Proof. By coordinate-independence of integrals, it suffices to prove the theorem for $M = U \subset \mathbb{R}_+^m$ and ω with compact support in U . We can write

$$\omega = \sum_{i=1}^{m-1} \omega_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^m, \quad (10.4.8)$$

where the hat again indicates exclusion. Then

$$d\omega = \sum_{i=1}^m (-1)^{i-1} \partial_i \omega_i dx^1 \wedge \cdots \wedge dx^m, \quad (10.4.9)$$

and

$$\int_U d\omega = \sum_{i=1}^m (-1)^{i-1} \int_U \partial_i \omega_i dx^1 \cdots dx^m. \quad (10.4.10)$$

If $\partial U = \emptyset$, the answer is zero by compactness of support. If $\partial U \neq \emptyset$, then

$$\int_{\mathbb{R}^{m-1}} \left(\int_0^\infty \partial_m \omega_m dx^n \right) dx^1 \cdots dx^m = - \int_{\mathbb{R}^{m-1}} \omega_m(x^1, \dots, x^{m-1}, 0) dx^1 \cdots dx^{m-1}, \quad (10.4.11)$$

and therefore

$$\int_U d\omega = (-1)^m \int_{\mathbb{R}^{m-1}} \omega_m(x^1, \dots, x^{m-1}, 0) dx^1 \cdots dx^{m-1}. \quad (10.4.12)$$

On the other hand, $\int_{\partial U} \omega = \int_{\partial \mathbb{R}_+^m} \omega = \int_{\partial \mathbb{R}_+^m} \omega(x^1, \dots, x^{m-1}, 0) dx^1 \cdots dx^{m-1}$ and the outward normal on ∂U is $-e_m = (0, \dots, 0, -1)$, therefore the induced orientation is $\text{sign}(-e_m, e_1, \dots, e_{m-1}) = (-1)^m$, so

$$\int_{\partial U} i^*\omega = (-1)^m \int_{\mathbb{R}^{m-1}} \omega_m(x^1, \dots, x^{m-1}, 0) dx^1 \cdots dx^{m-1} = \int_U d\omega. \quad (10.4.13)$$

Corollary 10.4.6. If M^m is compact and $\partial M = \emptyset$, then for any $\alpha \in \Omega^p(M)$, $\beta \in \Omega^{m-p}(M)$, and $X \in \mathfrak{X}(M)$ the following version of integration by parts holds:

$$\int_M (\mathcal{L}_X \alpha) \wedge \beta = - \int_M \alpha \wedge \mathcal{L}_X \beta. \quad (10.4.14)$$

Proof. We use the fact that \mathcal{L}_X is a derivation and Cartan's magic formula: $\int_M (\mathcal{L}_X \alpha) \wedge \beta + \alpha \wedge \mathcal{L}_X \beta = \int_M \mathcal{L}_X(\alpha \wedge \beta) = \int_M (d\mathbf{i}_X + i_X d)(\alpha \wedge \beta) = \int_M i_X d(\alpha \wedge \beta)$. But the exterior derivative of any top form is zero, so $d(\alpha \wedge \beta) = 0$. ■

Corollary 10.4.7. Under the conditions of the last Corollary, if μ is a volume form on M and $X \in \mathfrak{X}(M)$ is volume-preserving, i.e. $\operatorname{div}_\mu X = 0$, then for any $f, g \in C^\infty(M)$

$$\int_M (\mathcal{L}_X f) \cdot g \cdot \mu = - \int_M f \cdot (\mathcal{L}_X g) \cdot \mu \quad (10.4.15)$$

Proof. Exercise. ■

10.5 Vector bundle-valued differential forms

Here we define differential forms that take values in a vector space, and more generally in a vector bundle.

Definition 10.5.1 (Vector-valued and twisted differential forms). Let $E \xrightarrow{\pi} M$ be a **VB** (in particular it can be a trivial bundle $M \times V$ with a vector space V). E -valued differential forms of degree p are defined as sections of $\wedge^p M \otimes E$:

$$\Omega^p(M, E) := \Gamma \left(\bigwedge^p M \otimes E \right). \quad (10.5.1)$$

If E is a line bundle, then we also say that the exterior bundle has been *twisted* by E , and its sections are called *twisted differential forms*.

For example, if we twist E by the line bundle of pseudo-scalar densities (i.e. scalars depending on local orientation), we obtain the bundle of pseudo-forms. Top degree pseudo-forms are called densities. Similarly, twisting by the determinant bundle gives “densitised forms”.

If $\{e_1, \dots, e_k\}$ is a local frame of E , then any $\omega \in \Omega^p(M, E)$ can be expanded as

$$\omega = \sum_{a=1}^k \sum_{1 \leq i_1 < \dots < i_p \leq n} \omega_{i_1 \dots i_p}^a dx^{i_1} \wedge \dots \wedge dx^{i_p} \otimes e_a = \sum_{a=1}^k \omega^a \otimes e_a, \quad \omega^a \in \Omega^p(U). \quad (10.5.2)$$

The wedge product $\wedge : \Omega^p(M, E_1) \otimes \Omega^q(M, E_2) \rightarrow \Omega^{p+q}(M, E_1 \otimes E_2)$ is defined as

$$(\omega \wedge \eta)(X_1, \dots, X_{p+q}) = \frac{1}{(p+q)!} \sum_{\pi \in S_{p+q}} (\text{sign } \pi) \omega(X_{\pi(1)}, \dots, X_{\pi(p)}) \otimes \eta(X_{\pi(p+1)}, \dots, X_{\pi(p+q)}), \quad (10.5.3)$$

or locally

$$\left(\sum_a \omega^a \otimes e_a \right) \wedge \left(\sum_b \eta^b \otimes e'_b \right) = \sum_{a,b} \omega^a \wedge \eta^b \otimes e_a \otimes e'_b. \quad (10.5.4)$$

Similarly, one can twist complex differential forms on a complex manifold by a complex vector bundle (the tensor products then being over complex numbers).

As for the exterior derivative, there is generally no way to define it for vector-valued forms. Suppose we attempt to define d on $\Omega^p(M, E)$ locally by a formula like $d(\omega^a \otimes e_a) = d\omega^a \otimes e_a$. Then for $f > 0 \in C^\infty(U)$, we would have

$$\begin{aligned} d(\omega^a \otimes e_a) &= d(f\omega^a \otimes f^{-1}e_a) = f d\omega^a \otimes f^{-1}e_a + (df) \wedge \omega^a \otimes f^{-1}e_a = \\ &= d\omega^a \otimes e_a + \underbrace{d(\ln f) \wedge \omega^a \otimes e_a}_{\neq 0}, \end{aligned} \quad (10.5.5)$$

which is not as much a contradiction as a statement that this definition of the derivative relies on a specific choice of local frames (more precisely, equivalence class of local frames) and thus is not natural. Moreover, to make d globally well-defined, we need to have a similar equality with f replaced by the transition functions of E . Then for the extra term $\propto d \ln \tau_{\alpha\beta}$ to vanish, we must be able to choose constant transition functions $\tau_{\alpha\beta}$, i.e. E needs to be a flat bundle. So, for differential forms valued in a flat bundle, we can define a consistent, albeit not natural, exterior derivative that has all the standard properties of d .

Example 10.5.2. If we're working with vector-valued forms, i.e. $\Omega^*(M, V)$ with V a vector space, then these are simply antisymmetric r -linear maps from $(TM)^r$ to V . The bundle of such r -forms can be identified with the tensor product of bundles $\bigwedge^r T^*M \otimes (M \times V)$ and the space of its sections is isomorphic to $\Omega^r(M, V) \cong \Omega^r(M) \otimes V$. If a basis $\{e_i\}$ of V is chosen, then every V -valued r -form α can be written as

$$\alpha = \sum_i \alpha^i \otimes e_i, \quad (10.5.6)$$

where $\alpha^i \in \Omega^r(M)$ are uniquely determined ordinary r -forms and e_i can be interpreted as either elements of V or as constant global sections of $M \times V$. As a consequence, with respect

to a local chart (U, x) the form α has the local representation

$$\alpha|_{\alpha} = \sum_{i,i_1,\dots,i_r} \alpha_{i_1,\dots,i_r}^i dx^{i_1} \wedge \cdots \wedge dx^{i_r} \otimes e_i. \quad (10.5.7)$$

A natural exterior product exists on $\Omega^*(M, V)$ only if V carries an additional structure of an algebra (such as matrix multiplication or a Lie bracket), in which case we can write (denoting the multiplication on V by \bullet)

$$\begin{aligned} (\eta_1 \wedge \eta_2)(X_1, \dots, X_{r_1+r_2}) &= \\ &= \frac{1}{r_1!r_2!} \sum_{\pi \in S_{r_1+r_2}} \text{sign}(\pi) \eta_1(X_{\pi(1)}, \dots, X_{\pi(r_1)}) \bullet \eta_2(X_{\pi(r_1+1)}, \dots, X_{\pi(r_1+r_2)}). \end{aligned} \quad (10.5.8)$$

Remark 10.5.3. If E is not flat, then we need to specify a covariant derivative ∇ on it (it acts on tensors by raising their covariant rank by one, see Remark 9.3.8), and then define the covariant exterior derivative using the following (graded) tensorial Leibniz rule

$$d_{\nabla}(\omega^a \otimes e_a) = d\omega^a \otimes e_a + (-1)^{\deg \omega} \omega^a \wedge \nabla e_a. \quad (10.5.9)$$

Then d_{∇} satisfies the usual Leibniz rule just like d , however now $d_{\nabla}^2 \neq 0$! In fact, d_{∇}^2 is an End E -valued 2-form, called the *curvature form* $d_{\nabla}^2 = F_{\nabla} \in \Omega^2(M, \text{End}(E))$. If there exists a connection ∇ on E such that $F_{\nabla} \equiv 0$, then E is in fact a flat bundle, and we go back to the previous case (indeed, a choice of an equivalence class of local frames related by constant transition functions is equivalent to a choice of a flat connection). We will explore the details of this in Part VI.

10.6 Frobenius theorem II

The notion of Pfaffian system is dual to the notion of distribution that we used for the original Frobenius theorem. It plays an important role in the theory of first order partial differential equations and in many physical applications, notably in thermodynamics and in mechanical systems with constraints.

For a distribution $\Delta \subset T^*M$, let $\Omega_{\Delta}^1(M)$ denote the set of 1-forms on M with values in Δ .

Definition 10.6.1 (Pfaffian system). A Pfaffian system on M is a distribution $\Delta \subset T^*M$ such that for every $\beta \in \Delta_m$ there exists $\alpha \in \Omega_{\Delta}^1(M)$ such that $\alpha_m = \beta$.

Definition 10.6.2 (Annihilator). Given a vector space V , the annihilator of a subspace $W < V$ is the subspace $W^0 < V^*$ defined as

$$W^0 = \{\alpha \in V^* : \alpha|_W = 0\}. \quad (10.6.1)$$

Let $E \xrightarrow{\pi} M$ be a **VB** and $F \subset E$ a subbundle. Then

$$F^0 = \bigcup_{m \in M} F_m^0 \quad (10.6.2)$$

is a subbundle of E^* , called the annihilator of F in E .

Remark 10.6.3. A Pfaffian system Δ is regular of rank r iff it is a rank r vector subbundle of T^*M . In this case it admits local frames built from local 1-forms on M . Let $\{\theta^i\}_{i=1}^r$ be such a local frame over $U \subset M$. If the local r -form $\theta = \theta^1 \wedge \cdots \wedge \theta^r$ is closed, the restriction of Δ to U coincides with the annihilator of the *characteristic distribution* of θ (its fibers are $D_m^\theta = \ker \theta_m \cap \ker(d\theta)_m$).

Definition 10.6.4 (Integral manifold). Let Δ be a Pfaffian system on M . A connected immersed submanifold (N, ψ) of M is called an integral manifold of Δ through $m \in M$ if $m \in \psi(N)$ and

$$T_p\psi(T_p N) = \Delta_{\psi(p)}^0 \quad (10.6.3)$$

for all $p \in N$, where $\Delta_{\psi(p)}^0 \subset T_{\psi(p)}M$ denotes the annihilator of $\Delta_{\psi(p)}$. The Pfaffian system Δ is said to be integrable if for every $m \in M$ there exists an integral manifold of Δ through m .

Remark 10.6.5. Assume that Δ is regular, i.e. a subbundle of T^*M . Then its annihilator $\Delta^0 \subset TM$ is a regular distribution of complementary rank. Clearly a regular Pfaffian system is integrable iff so is the distribution Δ^0 . In this case, the integral manifolds of Δ and Δ^0 coincide. Similarly, if $D \subset TM$ is a regular distribution, then the annihilator D^0 is a regular Pfaffian system.

Example 10.6.6.

1. Every set of 1-forms $A \subset \Omega^1(M)$ spans a Pfaffian system, and conversely every Pfaffian system is generated in this way. Note that this does not imply that Δ is trivial.
2. The Pfaffian system on \mathbb{R}^2 spanned by the 1-forms dx and xdy is singular because it has rank 1 on the y -axis and rank 2 outside. Similarly, the Pfaffian system spanned by dx and ydy is singular. The first system is integrable with integral manifolds being the y -axis and the single points outside. In contrast, the second system is not integrable.

Pfaffian systems naturally arise in the theory of PDE's. Consider the first order system

$$\partial_i y^j = f_i^j(x^1, \dots, x^p, y^1, \dots, y^{n-p}), \quad 1 \leq i \leq p, 1 \leq j \leq n-p, \quad (10.6.4)$$

in the unknowns $y^j \in C^\infty(\mathbb{R}^p)$, with $f_i^j \in C^\infty(\mathbb{R}^n)$. Since the graph of the combined map $y = (y^1, \dots, y^{n-p}) : \mathbb{R}^p \rightarrow \mathbb{R}^{n-p}$ is a p -dimensional embedded submanifold in \mathbb{R}^n and

since y solves (10.6.4) iff

$$dy^j - \sum_{i=1}^p f_i^j(x^1, \dots, x^p, y^1, \dots, y^{n-p}) dx^i = 0, \quad 1 \leq j \leq n-p, \quad (10.6.5)$$

solutions of (10.6.4) can be interpreted geometrically as integral manifolds of the Pfaffian system Δ on \mathbb{R}^n , spanned by the 1-forms

$$\theta^j = dx^{p+j} - \sum_{i=1}^p f_i^j(x^1, \dots, x^n) dx^i, \quad 1 \leq j \leq n-p. \quad (10.6.6)$$

If the θ^j are linearly independent, Δ is regular. Then Δ is integrable iff so is the annihilator $\Delta^0 = \ker \theta^1 \cap \dots \cap \ker \theta^{n-p}$. In the special case $p = n-1$, we obtain the following simple solvability criterion.

Proposition 10.6.7 ([16, Prop. 4.7.6]). *A regular Pfaffian system Δ of rank 1, spanned by $\theta \in \Omega^1(M)$, is integrable iff $d\theta(X_1, X_2) = 0$ for all $X_1, X_2 \in \mathfrak{X}^{\Delta^0}(M)$.*

Proof. Proposition 10.2.10 implies $d\theta(X_1, X_2) = -\theta([X_1, X_2])$. Since $\Delta^0 = \ker \theta$, the Frobenius Theorem 9.5.4 yields the assertion. ■

Now, we are going to derive an integrability criterion for regular Pfaffian systems in the general case. For a Pfaffian system Δ , let $\Omega_\Delta^*(M)$ denote the subspace of $\Omega^*(M)$ spanned by the forms $\alpha \in \Omega^r(M)$, $r \geq 1$, which satisfy $\alpha_m(X_1, \dots, X_r) = 0$ for all $X_1, \dots, X_r \in \Delta_m^0$ and $m \in M$. $\Omega_\Delta^*(M)$ is a two-sided ideal in the associative algebra $\Omega^*(M)$.

Definition 10.6.8 (Differential ideal). An ideal $J \subset \Omega^*(M)$ is called a differential ideal if $dJ \subset J$.

Theorem 10.6.9 (Frobenius Theorem for Pfaffian systems [16, Thm. 4.7.8]). *A regular Pfaffian system Δ on M is integrable iff $\Omega_\Delta^*(M)$ is a differential ideal.*

Proof. By Frobenius Theorem 9.5.4, we must show that $d\Omega_\Delta^*(M) \subset \Omega_\Delta^*(M)$ is a necessary and sufficient condition for Δ^0 to be involutive. That it is necessary follows from Proposition 10.2.10. To see that it is sufficient, let r denote the rank of Δ and let $m \in M$. Choose $\theta^1, \dots, \theta^r \in \Omega_\Delta^1(M)$ such that $\{\theta_m^i\}_{i=1}^r$ is a basis in Δ_m . According to Proposition 10.2.10, for $X, Y \in \mathfrak{X}^{\Delta^0}(M)$,

$$(d\theta^j(X, Y))(m) = -\theta_m^i([X, Y]_m), \quad 1 \leq i \leq r. \quad (10.6.7)$$

If $d\Omega_\Delta^*(M) \subset \Omega_\Delta^*(M)$, the left hand side vanishes, hence $[X, Y]_m \in \Delta_m^0$. Since m was arbitrary, it follows that Δ^0 is involutive. ■

Next, we derive local criteria for a regular Pfaffian system to be integrable. We need

Lemma 10.6.10 ([16, Lem. 4.7.9]). *Let Δ be a regular Pfaffian system of rank r on M , let $m \in M$ and let $\{\theta^i\}_{i=1}^r$ be a local frame in Δ at m . Then m admits an open neighborhood U such that for every k -form α in $\Omega_{\Delta}^*(M)$ there exist local $(k-1)$ -forms β_1, \dots, β_r over U such that*

$$\alpha|_U = \sum_{j=1}^r \theta^j \wedge \beta_j. \quad (10.6.8)$$

Proof. On some neighborhood U of m , the local frame $\{\theta^i\}_{i=1}^r$ in Δ can be complemented by local 1-forms $\theta^{r+1}, \dots, \theta^n$ to a local frame in T^*M at m . Let $\{X_1, \dots, X_n\}$ be the dual local frame in TM . Then, $\{X_{r+1}, \dots, X_n\}$ is a local frame in Δ^0 . Expand

$$\alpha|_U = \sum_{i_1 < \dots < i_k} \alpha_{i_1 \dots i_k} \theta^{i_1} \wedge \dots \wedge \theta^{i_k}. \quad (10.6.9)$$

Since $\alpha \in \Omega_{\Delta}^*(M)$, we have that $\alpha_{i_1 \dots i_k} = \alpha|_U(X_{i_1}, \dots, X_{i_k}) = 0$ whenever $r < i_1$. Hence $\alpha|_U = \sum_{j=1}^r \theta^j \wedge \beta_j$ with $\beta_j = \sum_{j < i_2 < \dots < i_k} \alpha_{j i_2 \dots i_k} \theta^{i_2} \wedge \dots \wedge \theta^{i_k}$. ■

Proposition 10.6.11 ([16, Prop. 4.7.10]). *Let Δ be a regular Pfaffian system of rank r on M . TFAE:*

1. Δ is integrable.
2. For every $m \in M$, there exist a local frame $\{\theta^i\}_{i=1}^r$ in Δ at m and local 1-forms γ_j^i , $i, j = 1, \dots, r$, such that for all $i = 1, \dots, r$ one has $d\theta^i = \sum_{j=1}^r \theta^j \wedge \gamma_j^i$.
3. For every $m \in M$, there exist a local frame $\{\theta^i\}_{i=1}^r$ in Δ at m such that for all $i = 1, \dots, r$ one has $d\theta^i \wedge \theta^1 \wedge \dots \wedge \theta^r = 0$.

Proof. 1 \Rightarrow 2: Choose a local frame $\{\theta^i\}_{i=1}^r$ in Δ over a neighborhood U of m . Since Δ is integrable, so is the Pfaffian system $\Delta_U = \Delta \cap T^*U$ on U . Hence, by Theorem 10.6.9, $d\theta^i \in \Omega_{\Delta_U}^*(M)$. Application of Lemma 10.6.10 to the manifold U and the Pfaffian system Δ_U yields the assertion.

2 \Rightarrow 1: According to Theorem 10.6.9, it suffices to show that $\Omega_{\Delta}^*(M)$ is a differential ideal. Let $\alpha \in \Omega_{\Delta}^*(M)$ be of degree k and let $m \in M$. Let $\{\theta^i\}_{i=1}^r$ be the local frame in Δ over a neighborhood U of m and let γ_j^i , $i, j = 1, \dots, r$, be the local 1-forms over U existing by assumptions. By Lemma 10.6.10, $\alpha|_U = \sum_{i=1}^r \theta^i \wedge \beta_i$ for appropriate local $(k-1)$ -forms β_i on U . Then

$$d(\alpha|_U) = \sum_{i,j=1}^r \theta^j \wedge \gamma_j^i \wedge \beta_i - \sum_{i=1}^r \theta^i \wedge d\beta_i \quad (10.6.10)$$

and hence $(d\alpha)_m(X_1, \dots, X_k) = 0$ for all $X_i \in \Delta_m^0$. Since m was arbitrary, we get $d\alpha \in \Omega_\Delta^*(M)$.

$2 \Rightarrow 3$ is obvious. $3 \Rightarrow 2$: We may complement the local frame $\{\theta^i\}_{i=1}^r$ in Δ by local 1-forms $\theta^{r+1}, \dots, \theta^n$ to a local frame in T^*M at m . Expanding $d\theta^i = \sum_{j < k} \alpha_{jk}^i \theta^j \wedge \theta^k$, we obtain

$$0 = \sum_{j < k} \alpha_{jk}^i \theta^j \wedge \theta^k \wedge \theta^1 \wedge \cdots \wedge \theta^r. \quad (10.6.11)$$

It follows that $\alpha_{jk}^i = 0$ for all $r < j$. Hence, $d\theta^i = \sum_{j=1}^r \theta^j \wedge \gamma_j^i$ with $\gamma_j^i = \sum_{j < k} \alpha_{jk}^i \theta^k$. ■

Example 10.6.12. Let $M = \mathbb{R}^3$ and let Δ be the Pfaffian system spanned by a nowhere-vanishing 1-form θ on M . According to Proposition 10.6.11(3), Δ is integrable iff $d\theta \wedge \theta = 0$. Writing $\theta = \theta_i dx^i$, we have

$$d\theta = (\partial_1 \theta_2 - \partial_2 \theta_1) dx^1 \wedge dx^2 + \text{c.p.}, \quad (10.6.12)$$

where “c.p.” stands for two more terms obtained by cyclic permutations of $(1, 2, 3)$. Thus $d\theta \wedge \theta = 0$ is equivalent to

$$(\partial_2 \theta_3 - \partial_3 \theta_2) \theta_1 + (\partial_3 \theta_1 - \partial_1 \theta_3) \theta_2 + (\partial_1 \theta_2 - \partial_2 \theta_1) \theta_3 = 0. \quad (10.6.13)$$

Thus, Δ is integrable iff the vector field $\sum_i \theta_i \partial_i$ is perpendicular to its curl.

Definition 10.6.13 (Integrating factor). Let $\theta \in \Omega^1(M)$. A nowhere-vanishing function $f \in C^\infty(M)$ is called an integrating factor for θ if the 1-form $f\theta$ is exact.

Proposition 10.6.14 ([16, Prop. 4.7.13]). *The Pfaffian system Δ spanned by a single nowhere-vanishing 1-form θ on M is integrable iff there exist local integrating factors for θ everywhere on M , i.e. for every $m \in M$ there is an open neighborhood $U \subset M$ and an integrating factor $f \in C^\infty(U)$ for $\theta|_U$. For every $g \in C^\infty(U)$ satisfying $f\theta|_U = dg$, the level set components of g are integral manifolds of Δ .*

Proof. First assume that a local integrating factor f exists on U . Then

$$0 = d(f\theta|_U) = df \wedge \theta|_U + f d\theta|_U, \quad (10.6.14)$$

hence $d\theta|_U = \theta|_U \wedge df/f$, and Proposition 10.6.11(2) yields that Δ is integrable. Conversely, if Δ is integrable, so is the regular distribution Δ^0 of rank $n-1$ on M . Hence, according, for example, to 9.5.5, for every $m \in M$ there exists a local chart (U, x) at m such that $\partial_i^{x^{n-1}}_{i=1}$ is a

local frame in Δ^0 at m . Then $\theta = hdx^n$ for some $h \in C^\infty(U)$. Since θ is nowhere-vanishing, so is h . Thus $f = h^{-1}$ is an integrating factor for $\theta|_U$.

Next, let $g \in C^\infty(U)$ be such that $f\theta|_U = dg$ and let Σ be a level set component of g . Since $dg = f\theta|_U$ is nowhere-vanishing, every point of U is regular for g . Thus Σ is an embedded submanifold of dimension $n - 1$ of U (and hence of M). Let $\iota : \Sigma \hookrightarrow M$ denote the natural inclusion. Due to

$$\iota^*\theta|_U = \frac{1}{f|_\Sigma}\iota^*(f\theta|_U) = \frac{1}{f|_\Sigma}\iota^*(dg) = \frac{1}{f|_\Sigma}d(g|_\Sigma) = 0, \quad (10.6.15)$$

one has $T_m\Sigma \subset \Delta_m^0$ for all $m \in \Sigma$. By counting dimensions we obtain equality. Hence Σ is an integral manifold of Δ . ■

Example 10.6.15 (Ideal gas). Consider a mole of an ideal gas described by the variables V (volume), T (temperature), and p (pressure), which fulfil the thermal equation of state

$$pV = RT. \quad (10.6.16)$$

Note that this defines a 2-dimensional surface $A \subset \mathbb{R}^3$, which for example can be parametrized by (V, T) with $V > 0$ and $T > 0$. A change of state is represented by a curve in A and the heat exchange δQ with the environment, by definition, is obtained by integrating the 1-form

$$\theta = c_VdT + p(V, T)dV \quad (10.6.17)$$

along this curve. Thus an adiabatic change of state (no heat exchange) corresponds to an integral manifold of the Pfaffian system Δ on A spanned by θ . By Proposition 10.6.11, Δ is integrable because $d\theta \wedge \theta = 0$ holds trivially on a 2-dimensional manifold. Then, by Proposition 10.6.14, there exist local integrating factors f for θ . To find them, it suffices to consider the condition $d(f\theta) = 0$. This yields

$$0 = c_V \frac{\partial f}{\partial V} dV \wedge dT + \frac{\partial f}{\partial T} pdT \wedge dV + f \frac{\partial p}{\partial T} dT \wedge dV, \quad (10.6.18)$$

and with (10.6.16) we conclude

$$c_V \frac{\partial f}{\partial V} - \frac{RT}{V} \frac{\partial f}{\partial T} = \frac{R}{V} f. \quad (10.6.19)$$

Any nowhere-vanishing function f which fulfills this differential equation is an integrating factor. Let us consider the following examples:

1. $f(T, V) = V^{R/c_V}$: a potential g of $f\theta$ can be obtained by a simple integration:

$$g(T, V) = c_V \left(TV^{R/c_V} - T_0 V_0^{R/c_V} \right). \quad (10.6.20)$$

The integral manifolds of the Pfaffian system Δ are given by the level set components of g in A ,

$$TV^{R/c_V} = \text{const.} \quad (10.6.21)$$

2. $f(T, V) = T^{-1}$: the potential g coincides with the entropy

$$S(T, V) = c_V \ln \frac{T}{T_0} + R \ln \frac{V}{V_0}, \quad (10.6.22)$$

and the integral manifolds of Δ provided by this integrating factor correspond to reversible processes. For the ideal gas, whose internal energy depends only on temperature, $dU = c_V dT$, we obtain

$$dS = \frac{c_V}{T} dT + \frac{R}{V} dV = \frac{1}{T} dU + \frac{p}{T} dV, \quad (10.6.23)$$

that is,

$$dU = T dS - p dV. \quad (10.6.24)$$

The integrability of the 1-form $\theta = c_V dT + p dV$ is equivalent to the second law of thermodynamics or, more precisely, it yields the mathematical foundation of this law for the ideal gas.

11 Lie Theory II: Geometric Structure

11.1 Invariant vector fields

In this section we show that the tangent bundle of a Lie group is trivial and has a natural set of global frames.

Proposition 11.1.1. *Any Lie group G is parallelizable, with the global trivialization of TG given by*

$$TG \mapsto G \times \mathfrak{g}, \quad (g, X_g) \mapsto \left(g, L_{g*}^{-1} X_g \right), \quad (11.1.1)$$

where $\mathfrak{g} = T_e G$ and $X_g \in T_g G$.

Proof. All we need to show is that L_{g*}^{-1} is a smooth map that is a fiberwise isomorphism of vector spaces. This is true by virtue of L_{g*}^{-1} being the derivative of a diffeomorphism. ■

Obviously there is a second natural trivialization by right translations.

Recall that a left-invariant vector field $X \in \mathfrak{X}_L(G)$ on a Lie group G is defined by the property $(L_g)_* X = X$ for all g . As we have shown before, all such vector fields are automatically smooth and complete (Propositon 9.2.2). We will now show that, under the Lie bracket of vector fields, they comprise a Lie algebra that is naturally isomorphic to our earlier constructions of the Lie algebra of G . When $X \in \mathfrak{g} = T_e G$, we will often use the notation X_R and X_L to denote the corresponding right- and left-invariant vector fields, respectively:

$$X_L(g) = (L_g)_* X, \quad X_R(g) = (R_g)_* X. \quad (11.1.2)$$

Construction III Recall that we denote by $\mathfrak{X}_L(G)$ the space of left-invariant vector fields on G . As we will now show, the Lie derivative endows this space with a structure of a Lie algebra that is isomorphic to our earlier constructions of Lie G .

Lemma 11.1.2. *If $X \in \mathfrak{X}_L(G)$, then the flow $F^X : G \times \mathbb{R} \rightarrow G$ is left-invariant, i.e.*

$$gF^X(h, t) = F^X(gh, t). \quad (11.1.3)$$

Moreover, if $\gamma(t)$ is an integral curve of X starting at $e \in G$, then

$$F^X(g, t) = R_{\gamma(t)}(g). \quad (11.1.4)$$

The analogous formula for right-invariant vector fields is $F^X(g, t) = L_{\gamma(t)}(g)$.

Proof. Since these flows are complete, to show their equality it is sufficient to prove that the vector fields generating them coincide:

$$\left. \frac{d}{dt} \right|_{t=0} (gF(h, t)) = (L_g)_* X_h = X_{gh} = \left. \frac{d}{dt} \right|_{t=0} F^X(gh, t). \quad (11.1.5)$$

To prove the second equation, it suffices to show that $R_{\gamma(t)}(g)$ is the integral curve of X starting at g , which follows trivially from the above:

$$R_{\gamma(t)(g)=L_g\gamma(t)} = L_g F_t^X(e) = F_t^X(g). \quad (11.1.6)$$

■

Lemma 11.1.3. *The space $\mathfrak{X}^G(G)$ is closed under the Lie bracket of vector fields (i.e. Lie derivatives).*

Proof. This follows from the naturalness of Lie derivatives (Proposition 9.3.7):

$$(L_g)_* \mathcal{L}_X Y = \mathcal{L}_{(L_g)_*} (L_g)_* Y = \mathcal{L}_X Y. \quad (11.1.7)$$

■

Theorem 11.1.4. *The evaluation map $E : (\mathfrak{X}^G(G), \mathcal{L}) \rightarrow (T_e G, \text{ad})$, $X \mapsto X_e$ is an isomorphism of Lie algebras.*

Proof. It is obvious that E is linear. Next we check that it intertwines the two Lie brackets. First note that for any $X \in \mathfrak{X}^G(G)$, we have $X_g = (L_g)_{*e} X_e$. Let $\gamma(t)$ be the integral curve of X starting at e . We compute, using the definitions of $\text{Ad}_{\gamma(t)}$ and ad_X :

$$\begin{aligned} E(\mathcal{L}_X Y) &= E \left(\left. \frac{d}{dt} \right|_{t=0} F_t^* Y \right) = \left. \frac{d}{dt} \right|_{t=0} R_{\gamma(t)}^* Y_{\gamma(t)} = \\ &= \left. \frac{d}{dt} \right|_{t=0} \left(R_{\gamma(t)}^{-1} \circ L_{\gamma(t)} \right)_{*e} Y_e = \left. \frac{d}{dt} \right|_{t=0} (\text{Ad}_{\gamma(t)} Y_e) = \text{ad}_{X_e} Y_e. \end{aligned} \quad (11.1.8)$$

Injectivity follows because each left-invariant vector field is uniquely determined by its value at e , and surjectivity follows similarly because each $X_e \in T_e G$ can be extended to a left-invariant field.

■

Corollary 11.1.5. *The space $\mathfrak{X}_R(G)$ of right-invariant vector fields, together with the Lie derivative, forms a Lie algebra that is anti-isomorphic to $\text{Lie } G$, with the anti-isomorphism*

given by the same evaluation map E :

$$E(\mathcal{L}_X Y) = -\text{ad}_{X_e} Y_e \quad (11.1.9)$$

for $X, Y \in \mathfrak{X}_R(G)$.

In summary, if $X, Y \in T_e G$, then we have the following formulas relating the Lie bracket on the Lie algebra to the Lie bracket of vector fields:

$$[X_L, Y_L] = ([X, Y])_L, \quad [X_R, Y_R] = -([X, Y])_R. \quad (11.1.10)$$

The following is an elementary form of the celebrated Maurer-Cartan equation which we will discuss in depth in the next section. On a fundamental level it demonstrates the existence of a canonical flat connection of G -principal fiber bundles, in particular on G itself. It is usually formulated in terms of differential forms, but here we provide an elementary version adapted from [36].

Proposition 11.1.6 (Maurer-Cartan equation). *Let $g : [0, 1]^2 \rightarrow G$ be a twice differentiable map $(t, s) \mapsto g(t, s)$, viewed as a homotopy of paths in G . Define the elements $\xi(t, s), \eta(t, s) \in \mathfrak{g}$ by the equations*

$$(R_g^{-1})_{*g} \partial_t g = \xi, \quad (R_g^{-1})_{*g} \partial_s g = \eta. \quad (11.1.11)$$

In simplified notation (which is literal e.g. in matrix groups) these read $\partial_t g = \xi g$, $\partial_s g = \eta g$. Then

$$\partial_t \eta - \partial_s \xi = [\xi, \eta]. \quad (11.1.12)$$

Proof. In the matrix case this is trivial because we can just differentiate the two equations with respect to s and t , respectively, then notice that $\partial_t \partial_s g = \partial_s \partial_t g$ and get the result.

In the general abstract case, first let us notice that these equations can be written as

$$\partial_t g = \xi_R(g), \quad \partial_s g = \eta_R(g). \quad (11.1.13)$$

where ξ_R and η_R are the corresponding right-invariant vector fields. Now let $f \in C^\infty(G)$ be a function and let us compute the mixed derivative of $f \circ g$ in two ways:

$$\partial_t \partial_s f(g(t, s)) = \partial_t (\eta_R \cdot f)(g(t, s)) = ((\partial_t \eta_R) \cdot f)(g(t, s)) + (\xi_R \cdot \eta_R \cdot f)(g(t, s)), \quad (11.1.14)$$

$$\partial_s \partial_t f(g(t, s)) = \partial_s (\xi_R \cdot f)(g(t, s)) = ((\partial_s \xi_R) \cdot f)(g(t, s)) + (\eta_R \cdot \xi_R \cdot f)(g(t, s)). \quad (11.1.15)$$

These two expressions must coincide because mixed derivatives of smooth functions on \mathbb{R}^2 don't depend on the order of partial derivatives. Since the equality holds for all f , we conclude the equality of differential operators $\partial_t \eta_R - \partial_s \xi_R = [\eta_R, \xi_R]$. The claim of the theorem follows by virtue of the homomorphism property of Lie derivatives, $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$, the natural isomorphism between Lie derivatives and vector fields, and the natural *anti-isomorphism* between right-invariant vector fields and the Lie algebra (which is why the sign of the commutator changes). ■

11.2 Invariant differential forms

Definition 11.2.1 (Coadjoint representation). The duals of the adjoint representations of G and \mathfrak{g} are called the coadjoint representations of G and \mathfrak{g} and are denoted by Ad^* and ad^* , respectively. By definition,

$$(\text{Ad}_g^* \xi)(Y) = \xi \left(\text{Ad}_g^{-1} Y \right), \quad g \in G, Y \in \mathfrak{g}, \xi \in \mathfrak{g}^*, \quad (11.2.1)$$

and

$$(\text{ad}_X^* \xi, Y) (Y) = -\xi ([X, Y]), \quad X, Y \in \mathfrak{g}, \xi \in \mathfrak{g}^*. \quad (11.2.2)$$

Once again, as in Proposition 11.5.6, we have

$$\ker \text{Ad}^* = C_G(G_0), \quad \ker \text{ad}^* = C(\mathfrak{g}). \quad (11.2.3)$$

Definition 11.2.2. A differential form ξ on a Lie group G is called left-invariant if $L_g^* \xi = \xi$ for all $g \in G$, and right-invariant if $R_g^* \xi = \xi$.

The set of left-invariant differential forms on G is denoted by $\Omega^*(G)^G$. Written pointwise, for 1-forms, the defining condition reads $\xi_{gh} \circ (L_g)_{*h} = \xi_b$ for all $g, h \in G$ or, equivalently,

$$\xi_a = \xi_e \circ (L_g^{-1})_{g*} \text{ for all } g \in G. \quad (11.2.4)$$

Proposition 11.2.3 ([16, Prop. 5.5.2]). Let G be a Lie group with Lie algebra \mathfrak{g} .

1. $\Omega^*(G)^G$ is a differential subalgebra of $\Omega^*(G)$ (i.e. is closed under exterior products and derivatives).
2. There exists natural isomorphisms of exterior algebras

$$\Omega^*(G)^G \cong \bigwedge T_e^* G \cong \bigwedge \mathfrak{g}^*. \quad (11.2.5)$$

The first one is given by $\xi \mapsto \xi_e$. The second one is induced by the isomorphism $\mathfrak{g} \cong T_e G$.

3. For all $\xi \in \Omega^r(G)^G$ and $X_1, \dots, X_r \in \mathfrak{g}$, seen as left-invariant vector fields, the function $\xi(X_1, \dots, X_r)$ on G is constant.

- Proof.*
1. $\Omega^*(G)^G$ is closed under wedge products because wedge products are natural. It is also closed under exterior derivatives because d commutes with pullbacks.
 2. It suffices to consider the map $\Omega^*(G)^G \rightarrow \bigwedge T_e^*G$, defined by $\xi \mapsto \xi_e$. This mapping is a homomorphism of algebras, which is injective by (11.2.4). To show that it is surjective, it is enough to prove that its image contains T_e^*G . Thus, let $\eta \in T_e^*G$. By fixing the second argument of the inverse left trivialization of T^*G ($(\xi_L^T)^{-1} : G \times T_e^*G \rightarrow T^*G$) to be η , one obtains a left-invariant 1-form ξ with $\xi_e = \eta$.
 3. This is a consequence of the definition of pullback for differential forms.

■

We will thus identify left-invariant differential r -forms with elements of $\bigwedge^r \mathfrak{g}^*$ without explicitly stating it.

Remark 11.2.4. 1. Using the general formula from Proposition 10.2.10, the exterior derivative of $\xi \in \bigwedge^r \mathfrak{g}^*$ is

$$d\xi(X_0, \dots, X_r) = \sum_{i < j} (-1)^{i+j} ([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_r), \quad (11.2.6)$$

where $X_0, \dots, X_r \in \mathfrak{g}$ and as usual hatted arguments are meant to be omitted. The right hand side may be taken as an intrinsic definition of a differential on the exterior algebra $\bigwedge \mathfrak{g}^*$, thus turning the natural isomorphic of algebras $\Omega^*(G)^G \cong \bigwedge \mathfrak{g}^*$ into an isomorphism of differential algebras.

2. For $\xi \in \mathfrak{g}^*$, (11.2.6) yields

$$d\xi(X, Y) = -\xi([X, Y]), \quad X, Y \in \mathfrak{g}. \quad (11.2.7)$$

Let $\{e_1, \dots, e_n\}$ be a basis in \mathfrak{g} , let c_{jk}^i be the corresponding structure constants and let $\{\alpha^1, \dots, \alpha^n\}$ be the dual basis in \mathfrak{g}^* . Then (11.2.7) is equivalent to

$$d\alpha^i = -\frac{1}{2} c_{jk}^i \alpha^j \wedge \alpha^k, \quad i = 1, \dots, n. \quad (11.2.8)$$

This can be seen by evaluating both sides on the basis elements e_i . This equation is actually another form of the Maurer-Cartan equation (11.1.12), which we will return to in a basis-free form later in this section.

3. In terms of left-invariant 1-forms, the inverse left and right trivializations $(\chi_L^T)^{-1}$ and $(\chi_R^T)^{-1}$ of T^*G read

$$(\chi_L^T)^{-1}(g, \xi) = \xi_a, \quad (\chi_R^T)^{-1}(g, \xi) = (\text{Ad}_g^* \xi)_g = \xi_g \circ (\text{Ad}_g)_{*g}. \quad (11.2.9)$$

Note that the 1-form $\text{Ad}_g^* \xi$ need not be left-invariant.

Proposition 11.2.3 implies, in particular, that the space of left-invariant n -forms (where $n = \dim G$) is one-dimensional and is spanned by $\alpha^1 \wedge \dots \wedge \alpha^n$ for any basis $\{\alpha^i\}$ in \mathfrak{g}^* . By left-invariance, every nonzero element of this space is a volume form.

Corollary 11.2.5. *One every Lie group there exists a left-invariant volume form v_G , also often written as $v_G(g) = dg$. This form is unique up to multiplication by a nonzero scalar.*

Definition 11.2.6 (Haar measure). The Lebesgue measure $d\mu_L$ associated with a left-invariant volume form on G is called a (left) Haar measure on G . Thus, Haar measures on G are left-invariant and unique up to a constant.

Remark 11.2.7. 1. If G is compact, one way to single out a unique left-invariant volume form is to require that the volume of G be equal to 1.
 2. Naturally, there is also a right-invariant volume form w_G , unique up to multiplication by a scalar, and a measure $d\mu_R$ associated with it. The natural question thus is how the two Haar measures are related. Clearly,

$$v_G(g) = \text{Ad}_g^* w_G(g) = \det(\text{Ad}_g) w_G(g). \quad (11.2.10)$$

Since the adjoint representation is a map $\text{Ad}_g : G \rightarrow \text{GL}(\mathfrak{g})$, its determinant is everywhere nonzero and is called the *modular quasicharacter*:

$$\Delta : G \rightarrow \mathbb{K}, \quad \Delta(g) = \det(\text{Ad}_g). \quad (11.2.11)$$

Therefore a *bi-invariant* measure on G exists if and only if $|\Delta| \equiv 1$, i.e. when Δ is a group character³¹. Such Lie groups are called *unimodular*.

If G is a connected real Lie group, then $\Delta > 0$ everywhere by continuity (since $\Delta(e) = 1$ always).

3. Compact Lie groups are unimodular. Indeed, since $\Delta : G \rightarrow \mathbb{K}^*$ is a homomorphism, $\Delta(G)$ must be a compact subgroup of \mathbb{K}^* , which means its image lies inside $\mathbb{S}^1 \subset \mathbb{C}$.

For compact Lie groups, where every smooth function is integrable with respect to any volume form, the existence of left-invariant volume forms provides the powerful tool of averaging over the group. This concept can be used, for example, to construct invariants of representations, e.g. an invariant scalar product.

Proposition 11.2.8 (Invariant scalar product [16, Prop. 5.5.6]). *Let G be a compact Lie group with a finite-dimensional representation $\rho : G \rightarrow \text{GL}(V)$ on a \mathbb{K} -vector space V . The vector space V admits a scalar product $\langle \cdot, \cdot \rangle$ such that*

$$\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle \text{ for all } g \in G, v, w \in V. \quad (11.2.12)$$

For the induced representation ρ_{*e} of \mathfrak{g} we have

$$\langle \rho_*(X)v, w \rangle + \langle v, \rho_*(X)w \rangle = 0 \text{ for all } X \in \mathfrak{g}, v, w \in V. \quad (11.2.13)$$

Thus every finite-dimensional representation of a compact Lie group may be assumed to be orthogonal (in case $\mathbb{K} = \mathbb{R}$) or unitary (in case $\mathbb{K} = \mathbb{C}, \mathbb{H}$).

Proof. Choose an arbitrary scalar product (\cdot, \cdot) on V , and a left-invariant volume form v_G on G , and define

$$\langle v, w \rangle = \int_G f_{v,w} v_G, \quad f_{v,w}(g) = (\rho(g^{-1})v, \rho(g^{-1})w). \quad (11.2.14)$$

The integral exists because G is compact. The defining properties of a scalar product carry over from (\cdot, \cdot) to $\langle \cdot, \cdot \rangle$. Let $g \in G$. A brief computation shows that

$$\langle \rho(g)v, \rho(g)w \rangle = \int_G (L_{g^{-1}}^* f_{v,w}) v_G = \int_G f_{v,w} v_G = \langle v, w \rangle, \quad (11.2.15)$$

as asserted. The formula for the Lie algebra follows by differentiation. ■

Corollary 11.2.9 ([16, Cor. 5.5.7]). *For a finite-dimensional \mathbb{K} -representation ρ of a compact Lie group G ,*

1. *The spectrum of $\rho(g)$ lies in $\mathbb{S}^1 \subset \mathbb{C}$, and the spectrum of $\rho_*(X)$ lies in $i\mathbb{R}$, for all $g \in G$ and $X \in \mathfrak{g}$.*
2. *In case $\mathbb{K} = \mathbb{C}$, $\rho(g)$ and $\rho_*(X)$ are diagonalizable for all $g \in G$ and $X \in \mathfrak{g}$.*

Proof. 1 follows from the properties of orthogonal and skew-symmetric matrices (which every representation of a compact group can be reduced to by the preceding Proposition 11.2.8). In the case $\mathbb{K} = \mathbb{C}$, $\rho(g)$ is unitary and $\rho_*(X)$ is skew-Hermitian, hence the second assertion follows from elementary linear algebra as well. ■

To conclude this section, we discuss left-invariant \mathfrak{g} -valued forms on G , that is, differential forms $\xi \in \Omega^*(G, \mathfrak{g})$ such that $L_g^* \xi = \xi$. Given a basis $\{e_i\}$ of \mathfrak{g} , every $\xi \in \Omega^r(G, \mathfrak{g})$ can be written as

$$\xi = \sum_{i=1}^n \xi^i \otimes e_i \quad (11.2.16)$$

with ordinary differential r -forms $\xi^i \in \Omega^r(G)$. The form ξ is left-invariant iff so are ξ^i . Since \mathfrak{g} is an algebra, $\Omega^*(G, \mathfrak{g})$ carries an exterior product $\Omega^{r_1}(G, \mathfrak{g}) \times \Omega^{r_2}(G, \mathfrak{g}) \rightarrow \Omega^{r_1+r_2}(G, \mathfrak{g})$,

given by formula (10.5.8) with \bullet replaced by the Lie bracket:

$$[\xi_1, \xi_2](Y_1, \dots, Y_{r_1+r_2}) = \frac{1}{r_1!r_2!} \sum_{\pi \in S_{r_1+r_2}} \text{sign}(\pi) [\xi_1(Y_{\pi(1)}, \dots, Y_{\pi(r_1)}), \xi_2(Y_{\pi(r_1+1)}, \dots, Y_{\pi(r_1+r_2)})] \quad (11.2.17)$$

Remark 11.2.10. The notation around Lie algebra-valued forms often gets very confusing because in matrix groups their wedge product can also be defined using the matrix product instead of the commutator, which changes the scalar factor in the front. For example, if A and B are two \mathfrak{gl}_n -valued 1-forms, then it is easy to see that

$$A \wedge B = \frac{1}{2}[A, B], \quad (11.2.18)$$

where the wedge product on the left is defined using (10.5.8) with \bullet standing for the matrix product, whereas on the right \bullet is replaced with the matrix commutator as in (11.2.17).

Sometimes the product $[A, B]$ is also denoted $[A \wedge B]$ to remind that these are differential forms. Unfortunately, in physics, where Lie groups are often assumed to consist of matrices, people often use the notation $A \wedge B$ to refer to $[A \wedge B]$, which is the most common source of confusion around various factors of 2. This gets especially confusing around expressions like $A \wedge A \wedge A$.

There holds the following analogue of Proposition 11.2.3.

Proposition 11.2.11 ([16, Prop. 5.5.10]). *Let G be a Lie group with Lie algebra \mathfrak{g} and let $\Omega_L^*(G, \mathfrak{g})^G$ denote the space of left-invariant \mathfrak{g} -valued differential forms on G .*

1. $\Omega^*(G, \mathfrak{g})^G$ is a differential subalgebra of $\Omega^*(G, \mathfrak{g})$.
2. There exist natural algebra isomorphisms

$$\Omega^*(G, \mathfrak{g})^G \cong \left(\bigwedge T^*G \right) \otimes \mathfrak{g} \cong \left(\bigwedge \mathfrak{g}^* \right) \otimes \mathfrak{g}, \quad (11.2.19)$$

given by $\xi \mapsto \xi_e$ and induced by the usual identification of the Lie algebra with $T_e G$, respectively.

3. For all $\xi \in \Omega^*(G, \mathfrak{g})^G$ and $X_1, \dots, X_r \in \mathfrak{g}$, the \mathfrak{g} -valued functions $\xi(X_1, \dots, X_r)$ on G is constant.

Proof. The arguments are completely analogous to those for ordinary left-invariant differential forms in Proposition 11.2.3, except for the surjectivity of the map $\xi \mapsto \xi_e$ because $\Omega^*(G, \mathfrak{g})$ need not be generated as an algebra by 1-forms. Thus, let $\eta \in \bigwedge^r T_e^*G \otimes \mathfrak{g}$. Choose a basis $\{e_i\}$ on \mathfrak{g} and write $\eta = \sum_i \eta^i \otimes e_i$ with $\eta^i \in \bigwedge^i T_e^*G$. According to Proposition 11.2.3(1),

$\eta^i = \xi_e^i$ for certain left-invariant $\xi^i \in \Omega^r(G)$. Then, $\xi = \sum_i \xi^i \otimes e_i$ is a left-invariant differential r -form with values in \mathfrak{g} satisfying $\xi_e = \eta$. ■

According to Proposition 11.2.11, left-invariant \mathfrak{g} -valued 1-forms can be identified with linear endomorphisms of \mathfrak{g} . One canonical endomorphism, the identity, gives rise to a very important 1-form.

Definition 11.2.12 (Maurer-Cartan form). The left-invariant \mathfrak{g} -valued 1-form on G corresponding to $\text{id}_{\mathfrak{g}}$ is called the (left) Maurer-Cartan form of G , denoted by Θ_G . By definition,

$$\Theta_G(X_L)(g) = X \text{ for all } X \in \mathfrak{g}, g \in G. \quad (11.2.20)$$

As a consequence, Θ_G assigns to a tangent vector Y at g the unique left-invariant vector field X such that $X_g = Y$. Moreover, given a basis $\{e_i\}$ on \mathfrak{g} and its dual basis $\{\alpha^i\}$ on \mathfrak{g}^* , identified with left-invariant 1-forms $\alpha^i \in \mathfrak{X}_L^*(G)$, Θ_G is given by

$$\Theta_G = \sum_i \alpha^i \otimes e_i. \quad (11.2.21)$$

Using (11.2.7), its exterior derivative is given by

$$d\Theta_G(X, Y) = -\langle \Theta_G, [X, Y] \rangle = -[X, Y] = -[\Theta_G(X), \Theta_G(Y)], \quad (11.2.22)$$

and hence

$$d\Theta_G + \frac{1}{2} [\Theta_G, \Theta_G] = 0. \quad (11.2.23)$$

This is the Maurer-Cartan equation (11.2.8) in a basis-free form. It is also equivalent to our original Maurer-Cartan equation (11.1.12), which is nothing but the pullback of this one along a map $g : [0, 1]^2 \rightarrow G$.

Using the multiplication map $m : G \times G \rightarrow G$, define the map $\theta : TG \times TG \rightarrow TG$ by the identity

$$\theta \circ (\pi_* \times \pi_*) = m_*, \quad (11.2.24)$$

where $\pi_* \times \pi_*$ is the natural isomorphism $(G \times G) \rightarrow TG \times TG$. It is easy to check that θ can be expressed as

$$\theta((g, \xi), (h, \eta)) = (gh, R_{h*}\xi + L_{g*}\eta). \quad (11.2.25)$$

Then another way to see that Θ_G is a smooth form is to represent it as the composition of smooth maps, using the zero section $\sigma_0 : G \rightarrow TG$:

$$\begin{array}{ccccccc} TG & \xrightarrow{\pi \times \text{id}} & G \times TG & \xrightarrow{i \times \text{id}} & G \times TG & \xrightarrow{\sigma_0 \times \text{id}} & TG \times TG \\ (g, \xi) & \mapsto & g \times (g, \xi) & \mapsto & g^{-1} \times (g, \xi) & \mapsto & (g^{-1}, 0) \times (g, \xi) \\ & & & & & & \mapsto (e, L_{g*}^{-1}\xi) \end{array}. \quad (11.2.26)$$

Remark 11.2.13. As mentioned in Remark 11.2.10, in matrix groups it is common to see the Maurer-Cartan equation written as

$$d\Theta_G + \Theta_G \wedge \Theta_G = 0, \quad (11.2.27)$$

and in physics literature there may be a factor of $\frac{1}{2}$ in front of the wedge product due to the implied commutator.

Also, since $\Theta_G(X_L) = X$ and $(X_L)_g = L_{g^{-1}e}X$, we also have $\Theta_G(X) = L_{g^{-1}g}^{-1}X_L(g)$ for any $g \in G$. Since left-invariant vector fields span all of TG , this means that on matrix groups we can write

$$(\Theta_G)_g = g^{-1}dg, \quad (11.2.28)$$

where dg stands for the identity tensor on TG (i.e. $dg(X_g) = X_g$ at any $g \in G$). This is the most common way of defining Θ_G in physics.

Remark 11.2.14. Of course, everything above similarly applies to right-invariant forms, and one can define a right-invariant Maurer-Cartan form Θ_G^R , which in matrix groups is simply dgg^{-1} .

Remark 11.2.15. As we have seen in Proposition 10.6.11, a Pfaffian system spanned by a nowhere-vanishing one-form θ (or equivalently the distribution $\ker \theta$) is integrable iff $d\theta = \xi \wedge \theta$ at least locally for some 1-form ξ . Then the Maurer-Cartan equation, written in the form (11.2.8), expresses the fact that Θ_G defines a canonical integrable Pfaffian system on G spanned by $n = \dim G$ left-invariant 1-forms (identified e.g. with any basis of \mathfrak{g}^*). Its only maximal integral manifold is G itself. As we will see, the Maurer-Cartan equation can be used to derive the unique real-analytic structure on any Lie group. This is one example of deep connections between integrability and certain differentiable structures.

Remark 11.2.16. The Maurer-Cartan form provides an easy way to write down a left-invariant volume form on a Lie group in coordinates. All one needs to do is compute the Maurer-Cartan form and then take its n -th power, where $n = \dim G$, to get a form $\omega = \Theta_G^n = \Theta_G \wedge \cdots \wedge \Theta_G$. This is still a \mathfrak{g} -valued form which must have the form $v_G \otimes X$ for some $X \in \mathfrak{g}$. Whatever basis one chooses on \mathfrak{g} , it will suffice to find any non-zero component of X in this expression to get a volume form on G . In practice this is even simpler: one computes the matrix of 1-forms $g^{-1}dg$ and then takes the wedge product of all independent components of this matrix (there are n of them).

Example 11.2.17. An illustrative example is provided by SO_3 . Here we will use the Euler factorization

$$R : (-\pi, \pi) \times (0, \pi) \times (-\pi, \pi) \rightarrow \mathrm{SO}_3, \quad R(\varphi, \theta, \psi) = R_z(\varphi)R_y(\theta)R_z(\psi), \quad (11.2.29)$$

where R_z and R_y are the standard rotation matrices about the z and y axes, respectively. The

map R induces a local chart on SO_3 (with the exclusion of a measure zero set covered by $\theta = 0$). We then compute

$$\begin{aligned} R^{-1}dR &= R_z(\psi)^{-1}R_y(\theta)^{-1} \left(R_z(\varphi)^{-1}dR_z(\varphi) \right) R_y(\theta)R_z(\psi) + \\ &\quad + R_z(\psi)^{-1} \left(R_y(\theta)^{-1}dR_y(\theta) \right) R_z(\psi) + R_z(\psi)^{-1}dR_z(\psi). \end{aligned} \quad (11.2.30)$$

Each of these terms is centered around the Maurer-Cartan for an SO_2 subgroup. This should be obvious, but it is also easy to check that the corresponding invariant volume forms are simply $d\varphi$, $d\theta$, and $d\psi$, respectively. Indeed, we have

$$R_z(\varphi)^{-1}dR_z(\varphi) = \begin{pmatrix} 0 & -d\varphi & 0 \\ d\varphi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = L_z d\varphi, \quad (11.2.31)$$

where L_z is the generator of rotations about the z axis, and analogous formulas hold for $R_z(\psi)$ and $R_y(\theta)$ (using L_y for the latter). We now perform all of the matrix multiplications and simplify:

$$R^{-1}dR = \begin{pmatrix} 0 & -\cos\theta d\varphi - d\psi & \sin\theta \sin\psi d\varphi + \cos\psi d\theta \\ * & 0 & \sin\theta \cos\psi d\varphi - \sin\psi d\theta \\ * & * & 0 \end{pmatrix}, \quad (11.2.32)$$

where the missing components are recovered by anti-symmetry. An invariant volume form on SO_3 is then obtained by taking a wedge product of all three independent components. Before doing so, we notice that the xz and yz components will already give a factor of $d\varphi \wedge d\theta$, so we only need to take the $d\psi$ term from the xy component. Thus we find

$$\begin{aligned} (\sin\theta \cos\psi d\varphi - \sin\psi d\theta) \wedge (\sin\theta \sin\psi d\varphi + \cos\psi d\theta) \wedge d\psi &= \\ &= \sin\theta d\varphi \wedge d\theta \wedge d\psi, \end{aligned} \quad (11.2.33)$$

and the corresponding Haar measure is simply $\sin\theta d\varphi d\theta d\psi$ (it is bi-invariant because SO_3 is compact). We can compute the volume of the group by evaluating the integral

$$\iiint \sin\theta d\varphi d\theta d\psi = (2\pi)^2 \int_0^\pi \sin\theta d\theta = 8\pi^2. \quad (11.2.34)$$

Thus the normalized invariant volume form is

$$v_{\mathrm{SO}_3} = \frac{1}{8\pi^2} \sin\theta d\varphi \wedge d\theta \wedge d\psi. \quad (11.2.35)$$

Example 11.2.18. Another example is the affine group of the real line $G = \text{Aff}_1(\mathbb{R})$, which is the space $\mathbb{R} \times (0, \infty)$ with the product

$$(a, b) \cdot (x, y) = (a + bx, by), \quad a, x \in \mathbb{R}, \quad b, y > 0, \quad (11.2.36)$$

and has a faithful representation by 2×2 real matrices of the form

$$g(a, b) = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix}. \quad (11.2.37)$$

We readily confirm that

$$g^{-1}dg = \begin{pmatrix} \frac{db}{b} & \frac{da}{b} \\ 0 & 0 \end{pmatrix}. \quad (11.2.38)$$

Therefore the form $v_G = b^{-2}da \wedge db$ is left-invariant, however it is not right-invariant because G is not unimodular. Indeed, the modular quasicharacter can be computed by acting on this form with a right translation:

$$R_{g(x,y)}^* \frac{da \wedge db}{b^2} = \frac{d(a + bx) \wedge d(by)}{(by)^2} = \frac{da \wedge db}{b^2y}, \quad (11.2.39)$$

hence $\Delta(x, y) = y^{-1}$. Thus the form $w_G = \Delta^{-1}v_G = b^{-1}da \wedge db$ is a right-invariant volume form on the affine group.

Exercise 11.1. As we already know, SU_2 consists of complex matrices of the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad |a|^2 + |b|^2 = 1. \quad (11.2.40)$$

Thus it is very easy to compute the Maurer-Cartan form

$$\begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} \begin{pmatrix} da & db \\ -db & d\bar{a} \end{pmatrix} = \begin{pmatrix} \bar{a}da + bdb\bar{b} & \bar{a}db - bd\bar{a} \\ \bar{b}da - adb\bar{b} & * \end{pmatrix} \quad (11.2.41)$$

and thus, after taking a wedge product of these three matrix components and using the fact that $ad\bar{a} + \bar{a}da + bdb\bar{b} + \bar{b}db = 0$ (the differential of the condition $|a|^2 + |b|^2 = 1$), we find that a bi-invariant volume form is given by

$$v_{SU_2} = \frac{da \wedge db \wedge d\bar{b}}{a}. \quad (11.2.42)$$

Example 11.2.19. $\mathrm{SL}_2(\mathbb{C})$ is not compact but is nevertheless unimodular (obviously). Its components $a, b, c, d \in \mathbb{C}$ satisfy $ad - bc = 1$, so it's a 6-dimensional real group. An invariant 3-form, similarly to SU_2 above, is given by

$$\xi = \frac{da \wedge db \wedge dc}{a}, \quad (11.2.43)$$

and to get an invariant volume form we take its square, $v_{\mathrm{SL}_2(\mathbb{C})} = i\xi \wedge \bar{\xi}$ (the factor of i makes it real-valued).

Example 11.2.20. On $\mathrm{GL}(n, \mathbb{R})$, since it is an open subset of \mathbb{R}^{n^2} , a (non-invariant) volume form is given by the standard volume $\omega_x = dx = \prod_{i,j=1}^n dx_{ij}$. An action by a left translation will produce a factor of the Jacobian $l(x) = \det(L_{x*})$, and a right translation will give $r(x) = \det(R_{x*}^{-1})$. The modular quasicharacter is then $\Delta(x) = \det \mathrm{Ad}_x = l(x)/r(x)$. For $\mathrm{GL}_n(\mathbb{R})$, it is clear that the Jacobian is simply $l(x) = (\det x)^n$ (because it stretches the volume of each column by a factor of $\det x$). Therefore a left-invariant volume form on $\mathrm{GL}_n(\mathbb{R})$ is given by

$$v_{\mathrm{GL}_n(\mathbb{R})}(x) = \frac{dx}{(\det x)^n}. \quad (11.2.44)$$

Exercise 11.2. The Euler angle parametrization of SU_2 is the map

$$U : (-\pi, \pi) \times (0, \pi) \times (-2\pi, 2\pi) \rightarrow \mathrm{SU}_2 \quad (11.2.45)$$

given by

$$U(\varphi, \theta, \psi) = \begin{pmatrix} e^{\frac{i}{2}(\psi+\varphi)} \cos \frac{\theta}{2} & -e^{\frac{i}{2}(\psi-\varphi)} \sin \frac{\theta}{2} \\ e^{-\frac{i}{2}(\psi-\varphi)} \sin \frac{\theta}{2} & e^{-\frac{i}{2}(\psi+\varphi)} \cos \frac{\theta}{2} \end{pmatrix}. \quad (11.2.46)$$

Write down the left-invariant volume form on SU_2 in these coordinates and normalize it so that the volume of the group is 1. It is easy to predict the result: it is the same expression as on SO_3 , except the normalizing factor will be $16\pi^2$ because the range of ψ is twice as large; this agrees with the fact that SU_2 is a double cover of SO_3 .

11.3 Existence of homomorphisms

The Maurer-Cartan form provides the infinitesimal information on the multiplication while the Maurer-Cartan equation gives the only local obstruction to its integrability. The explicit local formulation is contained in the following theorem.

Definition 11.3.1 (Logarithmic derivative of group-valued functions). Let G be a Lie group and M a smooth manifold. For a smooth map $f : M \rightarrow G$, its (left) logarithmic derivative

$\delta f : TM \rightarrow \mathfrak{g}$ is defined by

$$\delta f(\xi_m) = \Theta_G(f(m))(T_m f \cdot \xi_m) = L_{f(m)*}^{-1} T_m f = f^* \Theta_G(m), \quad \xi_m \in T_m M. \quad (11.3.1)$$

In particular, $\Theta_G = \delta \text{id}_G$.

Theorem 11.3.2 (Fundamental Theorem, local version [41, Thm. 6.1]). *Let G be a Lie group and \mathfrak{g} its Lie algebra, and let $\omega \in \Omega^1(M, \mathfrak{g})$ be a \mathfrak{g} -valued 1-form on a smooth manifold M . Then for any $x \in M$ there exists an open neighborhood $U \subset M$ of x and a function $f : U \rightarrow G$ such that $\delta f = \omega$ iff ω satisfies the structural equation*

$$d\omega(X, Y) + [\omega(X), \omega(Y)] = 0, \quad \forall X, Y \in \mathfrak{X}(M). \quad (11.3.2)$$

If M is connected and $f_1, f_2 : M \rightarrow G$ satisfy $\delta f_1 = \delta f_2$, then there is a unique element $c \in G$ (integration constant) such that $f_2(x) = c \cdot f_1(x)$ for all $x \in M$.

Proof. By explicit expansion using (11.2.26), one can verify that the Maurer-Cartan form behaves in the following way with respect to the multiplication and inversion maps:

$$(m^* \Theta_G)(\xi_g, \eta_h) = \text{Ad}_h^{-1} (\Theta_G(\xi_g) + \Theta_G(\eta_h)), \quad (11.3.3)$$

$$(i^* \Theta_G)(\xi_g) = -\text{Ad}_g(\Theta_G(\xi_g)). \quad (11.3.4)$$

Now we prove the last statement of the theorem. Define $h : M \rightarrow G$ by $h(x) = f_2(x)f_1(x)^{-1}$. We have to show that h is constant, for which it suffices to show that $h^* \Theta_G = 0$ since this implies that $\Theta_G \circ Th$ and thus Th is zero. By definition, we have $h = m \circ (\text{id}, i) \circ (f_2, f_1) \circ \Delta$, where $\Delta(x) = (x, x)$. Using the above formulas we thus compute that for $X \in T_x M$,

$$\begin{aligned} (h^* \Theta_G)(X) &= (m^* \Theta_G)(T_x f_2 \cdot X, T_x(i \circ f_1) \cdot X) = \\ &= \text{Ad}_{f_1(x)} (\Theta_G(T_x f_2 \cdot X)) + (i^* \Theta_G)(T_x f_1 \cdot X) = \\ &= \text{Ad}_{f_1(x)} (\delta f_2(x)(X) - \delta f_1(x)(X)) = 0. \end{aligned} \quad (11.3.5)$$

Concerning existence, $d(f^* \Theta_G)(X, Y) + [f^* \Theta_G(X), f^* \Theta_G(Y)]$ clearly vanishes by virtue of being the pullback of the Maurer-Cartan equation. Thus, it suffices to prove that for $\omega \in \Omega^1(M, \mathfrak{g})$ satisfying the equation we can find a function $f : M \rightarrow G$ such that $\omega = \delta f$. To do this, we (locally) construct the graph of f as a leaf of an integrable distribution on $M \times G$. Consider $\Omega = \pi_M^* \omega - \pi_G^* \Theta_G$, where π_M, π_G are the natural projections of the direct product $M \times G$. Identifying $T(M \times G)$ with $TM \times TG$, the kernel of Ω is given by the set of all (X, Y) such that $\omega(X) = \Theta_G(Y)$. Since Θ_G restricts to a linear isomorphism on each tangent space, there is a unique solution Y for this equation for any chosen X , so the

distribution $\ker \Omega$ is regular, its rank equals the dimension of M , and $T\pi_M$ restricts to a linear isomorphism on each fiber of $\ker \Omega$. To check involutivity, we note that by construction

$$d\Omega((X_1, Y_1), (X_2, Y_2)) = d\omega(X_1, X_2) - d\Theta_G(Y_1, Y_2) = -[\omega(X_1), \omega(X_2)] + [\Theta_G(Y_1), \Theta_G(Y_2)], \quad (11.3.6)$$

and this obviously vanishes if both (X_i, Y_i) lie in $\ker \Omega$.

By Proposition 10.6.7 this implies integrability of the distribution $\ker \Omega$. Given $x \in M$ and $g \in G$, there is a submanifold $N \subset M \times G$ containing (x, g) integral for $\ker \Omega$. We have observed that $T\pi_M$ restricts to an isomorphism on each fiber of $\ker \Omega$, so $\pi_M : N \rightarrow M$ is a local diffeomorphism. Hence, we can find a neighborhood $U \subset M$ of x and a local inverse $j : U \rightarrow N$ of this projection. Defining $f := \pi_G \circ j$ we obtain a smooth function $f : U \rightarrow G$ and for $X \in T_x M$ we get $(T_x j)(X) = (X, \Theta_G^{-1}(\omega(X)))$, and thus $\Theta_G((T_x f)(X)) = \omega(X)$, which means $\omega = f^* \Theta_G$. ■

Corollary 11.3.3 ([41, Cor. 5.3]). *Let G be a Lie group with Lie algebra. Then $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$ is the unique map $f : G \rightarrow \text{GL}(\mathfrak{g})$ satisfying $f(e) = e$ and $f^* \Theta_{\text{GL}(\mathfrak{g})} = \text{ad}_{\Theta_G}$. If $g : I \rightarrow G$ is a path starting at e , then $\text{Ad}_g : I \rightarrow \text{GL}(\mathfrak{g})$ is the development, starting at $e \in \text{GL}(\mathfrak{g})$, of $\text{ad}_{g^* \Theta_G}$ on $\text{GL}(\mathfrak{g})$.*

Proof. The first statement follows because Ad satisfies both conditions and is unique by Theorem 11.3.2. The second statement follows by

$$\text{Ad}_g^* \Theta_{\text{GL}(\mathfrak{g})} = g^* \text{Ad}_{\text{GL}(\mathfrak{g})}^* = g^* \text{ad}_{\Theta_G} = \text{ad}_{g^* \Theta_G}. \quad (11.3.7)$$

The local theorem immediately implies its own global version in the case when $M = [0, 1]$ because the one-dimensional Maurer-Cartan equation is always true, and the interval is compact, so only a finite number of local reconstructions of f are needed.

Definition 11.3.4 (Development of ω). Let M be a manifold and $\omega \in \Omega^1(M, \mathfrak{g})$ where $\mathfrak{g} = \text{Lie } G$ for some Lie group G . Given a smooth path $\gamma : [0, 1] \rightarrow M$, the development of ω along γ starting at $g \in G$ is the unique function $\tilde{\gamma} : [0, 1] \rightarrow G$ such that $\tilde{\gamma}(0) = g$ and $\tilde{\gamma}^* \Theta_G = \gamma^* \omega$.

Exercise 11.3. Show that the endpoint $\tilde{\gamma}(1)$ of the development of a smooth path is invariant under smooth reparametrizations of the path.

Theorem 11.3.5 ([41, Thm. 7.7]). *If $\gamma_0, \gamma_1 : [0, 1] \rightarrow M$ are two smooth homotopic (with fixed ends) paths in a manifold M and $\omega \in \Omega^1(M, \mathfrak{g})$ satisfies the structural equation $d\omega + \frac{1}{2}[\omega, \omega] = 0$, then its developments, starting at $g \in G$, along γ_0 and γ_1 have the same*

endpoints.

Proof. Suppose $\gamma_0(1) = \gamma_1(1) = q$. Let $h : I \times I \rightarrow M$ be the homotopy from γ_0 to γ_1 . Now $h^*\omega$ is a \mathfrak{g} -valued 1-form on $I \times I$ that satisfies the structural equation. By Theorem 11.3.2 each point of $I \times I$ lies in a neighborhood on which $h^*\omega$ is the logarithmic derivative of a map into G . By compactness of $I \times I$ these local maps stitch into a global map $H : I \times I \rightarrow G$ such that $H(0, 0) = g$ and $H^*\Theta_G = h^*\omega$. Since $h(1, s) = q$ is constant for $s \in I$, it follows that $h^*\omega = H^*\Theta_G$ vanishes along the right edge of $I \times I$, and hence H is also constant there. In particular, $H(1, 0) = H(1, 1)$ and so the developments of ω along γ_0 and γ_1 end at the same point. ■

Thus each such 1-form produces a well-defined map from the fundamental group $\pi_1(M, x_0)$ to G by $\Phi([\gamma]) = \tilde{\gamma}(1)$. We immediately observe that the concatenation of loops in M produces products of corresponding group elements in G : $\Phi(\gamma_1 * \gamma_2) = (\tilde{\gamma}_1 * \tilde{\gamma}_2)(1) = \Phi(\gamma_1)\tilde{\gamma}_2(1) = \Phi(\gamma_1)\Phi(\gamma_2)$. Thus this is a homomorphism.

Definition 11.3.6 (Monodromy representation). If $\omega \in \Omega^1(M, \mathfrak{g})$ satisfies the structural equation, then its developments along closed loops starting at $x_0 \in M$ determine a homomorphism $\Phi_\omega : \pi_1(M, x_0) \rightarrow G$, called the *monodromy representation* of ω . The image of Φ_ω denoted by $\Gamma \subset G$ is called the *period group*, or the monodromy group. Assuming M is path-connected, different choices of x_0 lead to conjugate monodromy groups.

Proposition 11.3.7. *If $f : M \rightarrow G$ is smooth, then the development of δf along any path $\gamma : I \rightarrow M$ starting at $g = f(\gamma(0))$ ends at $f(\gamma(1))$. In particular, the monodromy representation of δf is trivial.*

Proof. Let $\omega = \delta f$. Since $(f \circ \gamma)^*\Theta_G = \gamma^*f^*\Theta_G = \gamma^*\omega$, the development of ω along γ is just $f \circ \gamma$. ■

Theorem 11.3.8 (Fundamental Theorem, global version [41, Thm. 7.14]). *Let G be a Lie group with Lie algebra \mathfrak{g} . Let M be a smooth connected manifold, let $x_0 \in M$, and let $\omega \in \Omega^1(M, \mathfrak{g})$. Then there exists a smooth map $f : M \rightarrow G$ such that $\delta f = \omega$ iff $d\omega + \frac{1}{2}[\omega, \omega] = 0$ and the monodromy representation $\Phi_\omega : \pi_1(M, x_0) \rightarrow G$ is trivial. Moreover, if there are two solutions f_1 and f_2 , then $f_1 \equiv c \cdot f_2$ for some $c \in G$.*

Proof. The uniqueness part was already shown in the local theorem 11.3.2. The forward implication was shown in Proposition 11.3.7.

Let us verify the backward implication. Define $f(x)$ to be the endpoint of the development of ω along any path from x_0 to x . This definition makes f well-defined under homotopies of

the path by Theorem 11.3.5 and under general changes in the path by the very definition of monodromy. We need to check only that $\omega = \delta f$.

Note that $f(x)$ can be obtained in two steps as follows. If $y \in M$, choose two paths $x_0 \rightsquigarrow y \rightsquigarrow x$. Then if the development of the first path starts at e and ends at g_0 , while the development of the second starts at e and ends at g , then the development of the composite path will start at e and end at $g_0 g$.

On the other hand, the local existence theorem guarantees that there is a connected neighborhood U of y and a smooth map $f_U : U \rightarrow G$ such that $\delta f_U = \omega$. After a left translation of f_U by some element of G we may assume that $f_U(y) = f(y)$. Now the development of ω starting at $f(y)$ along any curve $y \rightsquigarrow x$ ends at $f_U(x)$.

We conclude that $f_U(x) = f(x)$ for all $x \in U$, thus $\delta f = \delta f_U = \omega$. ■

Remark 11.3.9. The reason the above theorem is called fundamental is because in the case $M = G = \mathbb{R}$ it becomes the Fundamental Theorem of Calculus. If M is an open subset of \mathbb{R}^n and $G = \mathbb{R}^n$, then it states that a vector field $v^i(x^1, \dots, x^n)$ has a potential function iff $\partial_j v^i - \partial_i v^j = 0$ for all $1 \leq i \leq j \leq n$.

Theorem 11.3.10 (Integrability Theorem [41, Prop. 7.15]). *Let G and H be Lie groups with G simply connected. Then for every homomorphism of Lie algebras $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ there is a unique homomorphism of Lie groups $\Phi : G \rightarrow H$ such that $\Phi_{*e} = \phi$.*

Proof. We give a simple proof based on developments of 1-forms. We will also provide another instructive proof in Section 12.6 after we cover the exponential map. For yet another proof via group actions, see [7, Thm. 20.19].

Consider $\omega = \phi \circ \Theta_G \in \Omega^1(G, \mathfrak{h})$. This 1-form satisfies the structural equation because ϕ is linear and a Lie algebra homomorphism. Thus there is a unique map $\Phi : (G, e) \rightarrow (H, e)$ satisfying $\Phi^* \Theta_H = \omega$. We claim that Φ is a homomorphism. To see this, consider $f : (G, e) \rightarrow (H, e)$ defined by $f(g) = \Phi(g_0)^{-1} \Phi(g_0 g)$ for some fixed g_0 . Now $f = L_{\Phi(g_0)}^{-1} \circ \Phi \circ L_{g_0}$, and so

$$f^* \Theta_H = L_{g_0}^* \Phi^* L_{\Phi(g_0)}^{*-1} \Theta_H = L_{g_0}^* \Phi^* \Theta_H = L_{g_0}^* \omega = \omega. \quad (11.3.8)$$

Since f and Φ agree at the identity and have the same logarithmic derivative, it follows that $f = \Phi$ and hence $\Phi(g) = \Phi(g_0)^{-1} \Phi(g_0 g)$, that is, $\Phi(g_0) \Phi(g) = \Phi(g_0 g)$ for all $g, g_0 \in G$. ■

Remark 11.3.11. Note that uniqueness holds as long as G is connected, and it doesn't need to be simply connected. This is why instead of this single theorem we sometimes speak of two separate principles of Lie theory, the first one claiming uniqueness, and the second

claiming existence:

First Principle Assuming G is connected, every Lie group homomorphism $\Phi : G \rightarrow H$ is uniquely determined by its differential at the identity. (Note that the existence of Φ is assumed).

Second Principle Assuming G is simply connected, for any homomorphism of Lie algebras $\phi : \text{Lie } G \rightarrow \text{Lie } H$ there is a unique Lie group homomorphism $G \rightarrow H$ whose differential at the identity is ϕ .

Corollary 11.3.12. *Let G be a connected and simply connected Lie group. For every representation $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ of the Lie algebra $\mathfrak{g} = \text{Lie } G$ there is a unique representation $R : G \rightarrow \text{GL}(V)$ such that $R_{*e} = \rho$.*

Example 11.3.13. The Lie group SU_2 , being homeomorphic to the sphere \mathbb{S}^3 , is simply-connected. Therefore every homomorphism $\phi : \mathfrak{su}_2 \rightarrow \mathfrak{g}$ into any other Lie algebra \mathfrak{g} of some Lie group G determines a unique homomorphism $\Phi : SU_2 \rightarrow G$.

On the other hand, from Example 6.1.10, the homomorphism $\text{Ad} : SU_2 \rightarrow SO_3(\mathbb{R})$ is a double covering map since $\ker \text{Ad} = \{\pm I\}$, thus $SO_3(\mathbb{R})$ doesn't satisfy the condition on the domain group for the same theorem, and it must be possible to find a Lie algebra homomorphism $\psi : \mathfrak{so}_3(\mathbb{R}) \rightarrow \mathfrak{g}$ that doesn't extend to a group homomorphism $\Psi : SO_3(\mathbb{R}) \rightarrow G$. Namely, the group homomorphism Ψ exists if and only if

$$\exp \psi(A_x(2\pi)) = e_G, \text{ where } A_x(2\pi) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2\pi \\ 0 & 2\pi & 0 \end{pmatrix}. \quad (11.3.9)$$

To see this, apply the previous statement to $\phi = \psi \circ \text{ad}$, where we view the adjoint as a map $\text{ad} : \mathfrak{su}_2 \rightarrow \mathfrak{so}_3(\mathbb{R})$. From this homomorphism, Ψ is determined by $\Phi = \Psi \circ \text{Ad}$; therefore such a homomorphism of Lie groups Ψ exists iff $\Phi(-I) = e_G$, which translates to $\Psi \circ \text{Ad}(-I) = e_G$. Since $\Phi \circ \exp = \exp \circ \phi$ and

$$\exp(i\sigma_1) = \exp \begin{pmatrix} 0 & i\pi \\ i\pi & 0 \end{pmatrix} = -I, \quad (11.3.10)$$

the condition that $\Phi(-I) = e$ becomes

$$e = \Phi \circ \exp(i\sigma_1) = \exp \circ \phi(i\sigma_1) = \exp \circ \psi(\text{ad}_{i\sigma_1}) = \exp \circ \psi(R_x(2\pi)), \quad (11.3.11)$$

where at the end we used formula (6.10.36).

This picture generalizes to all Lie groups. For any group H one can specify a finite set of $X_i \in \mathfrak{h}$ such that any homomorphism of Lie algebras $\phi : \mathfrak{h} \rightarrow \mathfrak{g}$ generates a homomorphism of Lie groups $\Phi : H \rightarrow G$ iff $\exp \phi(X_i) = e_G$ for all i .

The following proposition interprets the monodromy representation in terms of deck transformations of the universal covering acting on G .

Proposition 11.3.14 (Monodromy and coverings [41, Prop. 8.1]). *Let $\omega \in \Omega^1(M, \mathfrak{g})$ satisfy the structural equation. Let $\pi : \tilde{M} \rightarrow M$ be the universal covering map of M and let $\text{Aut}(\pi) \cong \pi_1(M, x_0)$ be its group of deck transformations. The Fundamental Theorem 11.3.8 applied to the pair $(\tilde{M}, \pi^*\omega)$ gives a unique map $f : \tilde{M} \rightarrow G$ with $f(x_0) = e$. Then there is a unique homomorphism $\tilde{\Phi} : \text{Aut}(\pi) \rightarrow G$ satisfying $f \circ A = \tilde{\Phi}(A) \cdot f$ for all $A \in \text{Aut}(\pi)$.*

Proof. Every deck transformation $A \in \text{Aut}(\pi)$ preserves the form $\pi^*\omega$ (since $A^*\pi^*\omega = (\pi \circ A)^*\omega = \pi^*\omega$), and so we have

$$(f \circ A)^*\Theta_G = A^*f^*\Theta_G = A^*\pi^*\omega = \pi^*\omega = f^*\Theta_G. \quad (11.3.12)$$

By the Fundamental Theorem, $f \circ A$ and f differ by a left multiplication by a constant element $\tilde{\Phi}(A) \in G$, that is, $f \circ A = \tilde{\Phi}(A)f$. In particular, $f \circ A(p_0) = \tilde{\Phi}(A)f(p_0) = \tilde{\Phi}(A)$. To see that $\tilde{\Phi}$ is a homomorphism, note that

$$\tilde{\Phi}(AB) = f(AB(p_0)) = \tilde{\Phi}(A)f(B(p_0)) = \tilde{\Phi}(A)\tilde{\Phi}(B)f(p_0) = \tilde{\Phi}(A)\tilde{\Phi}(B). \quad (11.3.13)$$

■

The final result shows that in a certain sense the Maurer-Cartan form determines the group up to a covering map. (It is this description of a Lie group that will generalize to the notion of a Cartan connection on a principal bundle.)

Definition 11.3.15 (Complete 1-form). A vector-valued 1-form $\omega \in \Omega^1(M, V)$ is called complete if every vector field $X \in \mathfrak{X}(M)$ such that $\omega(X) = \text{const}$ is complete.

Theorem 11.3.16 ([41, Thm. 8.7]). *Let M be a connected smooth manifold and let \mathfrak{g} be a Lie algebra. Let $\omega \in \Omega^1(M, \mathfrak{g})$ satisfy the conditions*

- (i) $d\omega + \frac{1}{2}[\omega, \omega] = 0$,
- (ii) $\omega : TM \rightarrow \mathfrak{g}$ is an isomorphism on each fiber,
- (iii) ω is complete.

Then

- (a) *The universal cover of M , $\pi : G \rightarrow M$, has, for an arbitrary choice $e \in G$, the structure of a Lie group with identity element e and Lie algebra \mathfrak{g} whose Maurer-Cartan form is $\pi^*\omega$.*

(b) The period group $\Gamma = \Phi_\omega(\pi_1(M, x_0)) \subset G$ acts by left multiplication on G as the group of deck transformations for π .

Proof. First we discuss the case of simply connected M , so that $G = M$. This case involves five steps. First we construct a map $f : M \rightarrow \mathrm{GL}(\mathfrak{g})$ that will eventually be given by $f(g) = \mathrm{Ad}_g$. The second step constructs a map $m : M \times M \rightarrow M$ satisfying the conditions $m(e, e) = e$ and $m^* \omega = (\mathrm{pr}_2^* f)^{-1} \mathrm{pr}_1^* \omega + \mathrm{pr}_2^* \omega$ (see (11.3.3)). The third step shows that f is a homomorphism with respect to the multiplication m . The fourth shows that m is associative. The fifth step constructs a map $i : M \rightarrow M$ and shows that it is the inversion with respect to m .

Step 1. Corollary 11.3.3 shows that any map $f : M \rightarrow \mathrm{Aut}(\mathfrak{g})$ should have logarithmic derivative $\mathrm{ad}_\omega \in \mathfrak{gl}(\mathfrak{g})$ and satisfy $f(e) = e$. Thus the form ω will determine f , provided $\eta = \mathrm{ad}_\omega$ is integrable. But

$$d\eta + \frac{1}{2} [\eta, \eta] = \mathrm{ad}_{d\omega} + \frac{1}{2} [\mathrm{ad}_\omega, \mathrm{ad}_\omega] = \mathrm{ad}_{d\omega + \frac{1}{2} [\omega, \omega]} = 0, \quad (11.3.14)$$

and since M is simply connected, it follows from the Fundamental Theorem 11.3.8 that η is the logarithmic derivative of a unique map $f : M \rightarrow \mathrm{GL}(\mathfrak{g})$ satisfying $f(e) = e$. In particular, $df = f \circ \mathrm{ad}_\omega$, where the product on the right is the composition of elements in $\mathfrak{gl}(\mathfrak{g})$ as linear maps.

Step 2. We construct the multiplication map. First equation (11.3.3) shows that any multiplication map $m : M \times M \rightarrow M$ should have logarithmic derivative at $(x, y) \in M \times M$ given by (11.3.3):

$$\eta = \mathrm{pr}_1^*(\mathrm{Ad}_y^{-1} \omega) + \mathrm{pr}_2^* \omega = \mathrm{pr}_1^*(f(y)^{-1} \omega) + \mathrm{pr}_2^* \omega = (\mathrm{pr}_2^* f^{-1}) \mathrm{pr}_1^* \omega + \mathrm{pr}_2^* \omega. \quad (11.3.15)$$

Thus the form ω will determine m , provided η is integrable. Now one checks that $d\eta + \frac{1}{2} [\eta, \eta] = 0$ by a lengthy direct substitution and using the structural equation on ω , the equality $d(f^{-1}) = -\mathrm{ad}_\omega \circ f^{-1}$, and the symmetry of the bracket of Lie algebra-valued 1-forms. Since M is simply connected, the Fundamental Theorem shows that η is the log derivative of a unique map $m : M \times M \rightarrow M$ satisfying $m(e, e) = e$.

Step 3. We show that $f : M \rightarrow M$ is a homomorphism with respect to m in the sense that $f(m(x, y)) = f(x)f(y)$ for all $x, y \in M$. We shall the commutativity of the square consisting of m , m_{GL} , f , and $f \times f$:

$$m_{\mathrm{GL}} \circ (f \times f) = f \circ m. \quad (11.3.16)$$

It is clear that this square commutes at (e, e) . Since M is connected it suffices to show that the Maurer-Cartan form Θ_{GL} on $\mathrm{GL}(\mathfrak{g})$ pulls back to the same form on $M \times M$ along the two possible routes. Indeed, $\Theta_{\mathrm{GL}(\mathfrak{g})}$ trivializes the tangent bundle of $\mathrm{GL}(\mathfrak{g})$ and thus its

composition with the derivative of any map identifies that derivative, and hence the map itself, up to a constant. Thus we must show that

$$(f \circ m)^* \Theta_{\text{GL}} = (m \text{GL} \circ (f \times f))^* \Theta_{\text{GL}}. \quad (11.3.17)$$

Once again a lengthy substitution of $m^* \omega = \text{pr}_2^*(f^{-1}(\text{pr}_1^* \omega)) + \text{pr}_2^* \omega$ and $m_{\text{GL}}^* \Theta_{\text{GL}} = \text{pr}_2^* \text{Ad}_{\text{pr}_1^* \Theta_{\text{GL}}}^{-1} + \text{pr}_2^* \Theta_{\text{GL}}$ confirms that the above identity is true.

Step 4. Now we need to show that m is associative. This is equivalent to the commutativity of yet another square:

$$m \circ (m \times \text{id}) = m \circ (\text{id} \times m). \quad (11.3.18)$$

Again, this diagram clearly commutes at (e, e, e) , and since M is connected, it suffices to show that the form ω on M pulls back to the same form on $M \times M \times M$ along the two possible routes. Another lengthy substitution using the same identities verifies that this holds.

Step 5. Finally we construct an inversion map. Another trivial exercise is to check using (11.3.4) that any inversion map $i : M \rightarrow M$ must have log derivative given by

$$\eta_x(X) = -\text{Ad}_x \omega(X) = -(f\omega)(X). \quad (11.3.19)$$

Thus the form ω will determine i , provided we can show that $-f\omega$ is integrable. But since

$$d\eta + \frac{1}{2} [\eta, \eta] = -(df)\omega - f d\omega + \frac{1}{2} f[\omega, \omega] = -f \text{ad}_\omega \omega + f[\omega, \omega] = 0 \quad (11.3.20)$$

and M is simply connected, it follows from the Fundamental Theorem that η is the log derivative of a unique map $i : M \rightarrow M$ satisfying $i(e) = e$. However, we must still verify that i is the inverse map, i.e. that $m(i(x), x) = e$. This identity can be written as the composition

$$m \circ (i \times \text{id}) \circ \Delta = e \in M, \quad (11.3.21)$$

where $\Delta : M \rightarrow M \times M$ is the diagonal map $\Delta(x) = (x, x)$. Since M is connected, it suffices to show that the form ω on M pulls back to the same form on M along the two possible routes. Since it obviously pulls back to 0 via the right hand side, we must show that $(m \circ (i \times \text{id}) \circ \Delta)^* \omega = 0$. Once again, a lengthy substitution using the formula for $m^* \omega$ verifies this identity. Thus i is an inverse for μ .

The case $\pi_1(M) \neq 0$. Here the form ω on M pulls up to the form $\pi^* \omega$ on the universal cover G . Clearly it satisfies the same three conditions on G that ω does on M . Thus our proof applies to the pair $(G, \pi^* \omega)$ and we see that for any fixed choice $e \in G$ there is a Lie group structure on G for which e is the identity and $\pi^* \omega$ is the Maurer-Cartan form. By Proposition 11.3.14, there is a unique injective homomorphism $\tilde{\Phi} : \text{Aut}(\pi) \rightarrow G$ satisfying $\tilde{\Phi}(A)g = A(g)$ for all deck transformations $A \in \text{Aut}(\pi)$. Thus the group of deck transformations is identified with the period group, a discrete subgroup $\Gamma \subset G$. ■

- Remark 11.3.17.**
1. Dropping condition (iii) in Theorem 11.3.16 still leads to a manifold with a structure of a *local Lie group*.
 2. If we drop condition (i), we are left with what may be regarded as a “deformation” of a Lie group. For example we might consider altering the form ω , which does satisfy all three conditions, by $\omega_t = \omega + t\eta$ with η satisfying (ii). The new form ω_1 may also satisfy (iii), but it will not, in general, continue to satisfy (i), and so we will not even have a local Lie group.
 3. Except in special cases, we cannot weaken condition (ii) to ω_x being merely an injection for all $x \in M$ and still expect to obtain a discrete monodromy group.

Theorem 11.3.16 provides an easy proof of the existence and uniqueness of universal covering groups (Theorem 6.2.1).

Proof of Theorem 6.2.1. Let the universal covering map be $\pi : \tilde{G} \rightarrow G$. First note that if a Lie group structure on \tilde{G} exists, then the Maurer-Cartan form on \tilde{G} is $\pi^*\Theta_G$. Thus we are led to define $\omega = \pi^*\Theta_G$ on \tilde{G} . This form satisfies the structural equation by naturality of exterior derivatives and Lie brackets. Since π is a local diffeomorphism, not only is ω an isomorphism on each fiber, but also ω is complete, since integral curves for Θ_G on G will lift to integral curves for ω on \tilde{G} . By Theorem 11.3.16, given $\tilde{e} \in \pi^{-1}(e)$, \tilde{G} has a unique Lie group structure with \tilde{e} as the identity and for which ω is the Maurer-Cartan form. By Theorem 11.3.2 π is the unique map $(\tilde{G}, \tilde{e}) \rightarrow (G, e)$ pulling back ω from Θ_G , and then by Theorem 11.3.10 it must be a homomorphism. ■

11.4 One-parameter subgroups

Lemma 11.4.1. *If $\mathfrak{h} \subset T_e G$ is a Lie subalgebra of Lie G (identified with $T_e G$), then the subset $D_L = \bigcup_g (L_g)_{*e} \mathfrak{h} \subset TG = G \cdot \mathfrak{h}$ is a left-invariant involutive distribution. Similarly, $D_R = \mathfrak{h} \cdot G$ is a right-invariant involutive distribution. Moreover,*

$$\mathfrak{X}_L(G) = \Gamma^\infty(D_L), \quad \mathfrak{X}_R(G) = \Gamma^\infty(D_R). \quad (11.4.1)$$

Proof. It is sufficient to show this for the left-invariant case. $\bigcup_g (L_g)_{*e}$ is an isomorphism therefore D_{Lg} is of rank $\dim \mathfrak{h}$ and the distribution is obviously left-invariant. Any basis of left-invariant vector fields on G forms a smooth frame for D , so D is smooth. Moreover, D_L is involutive because such a frame is closed under the Lie bracket. ■

The Frobenius theorem now gives us the first the first “enhancement” result about Lie subgroups. Even though by definition they are only required to be immersed (which means that their topology can differ drastically from the subspace topology), the following theorem shows that they are in fact weakly embedded (a.k.a. initial).

Theorem 11.4.2 (Lie subgroups are weakly embedded [7, Thm. 19.25]). *Every Lie subgroup is an integral manifold of an involutive distribution, and therefore is a weakly embedded submanifold.*

Proof. Let $H \subset G$ be a Lie subgroup. Then $\text{Lie } H$ is canonically isomorphic to the Lie subalgebra $i_*(\text{Lie } H) \subset \mathcal{L}G$, where $i : H \hookrightarrow G$ is the inclusion. Let $D \subset TG$ be the left-invariant involutive distribution corresponding to \mathfrak{h} as in Lemma 11.4.1. It follows that $T_h H = D_h$, so H is an integral manifold of D and is weakly embedded by Theorem 9.5.7. ■

Recall from Theorem 3.9.14 and 3.9.15 that while an immersed submanifold can have its topology (and, in a uniquely corresponding way, smooth structure) adjusted in many ways while still remaining immersed, a weakly embedded submanifold has *only one* topology and smooth structure under which it's still immersed. This means that the topology of an immersed Lie subgroup is much more rigid than we originally assumed in the definition.

Remark 11.4.3. Note that this does not yet allow us to completely drop the assumption of any smooth structure in the definition of a Lie subgroup. A weakly embedded submanifold still needs to be immersed by definition, which presumes a predefined smooth structure on it. We have only shown that this smooth structure is unique. The existence of a smooth structure on an arbitrary path-connected subgroup of a Lie group is a much deeper result which we will demonstrate in Section 12.7

Theorem 11.4.4 ([7, Thm. 19.26]). *Let G be a Lie group and $\mathfrak{g} = \text{Lie } G$ its Lie algebra. If $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra, then there is a unique connected immersed Lie subgroup of G whose Lie algebra is \mathfrak{h} .*

Proof. Let $D \subset TG$ be the involutive distribution corresponding to \mathfrak{h} . We now use the global Frobenius theorem to introduce the foliation \mathcal{H} determined by D , and for any $g \in G$, let \mathcal{H}_g be the leaf containing g . Since D is left-invariant, it follows that left translations take leaves to leaves: $L_g(\mathcal{H}_{g'}) = \mathcal{H}_{gg'}$.

Define $H = \mathcal{H}_e$. We will show that H is the desired Lie subgroup. First, it is a subgroup because it is left-invariant: $hh' \in L_h(H) = \mathcal{H}_h = H$ for all $h, h' \in H$, and similarly $h^{-1} \in H$.

To show that H is a Lie group we need to show that the map $\mu(h, h') = hh'^{-1}$ is smooth as a map $H \times H \rightarrow H$. Since $H \times H$ is a submanifold of $G \times G$, it is immediate that $\mu : H \times H \rightarrow G$ is smooth. Since H is also weakly embedded by Theorem 9.5.7, μ is also smooth as a map into H .

The fact that H is a leaf of \mathcal{H} implies that the Lie algebra of H is \mathfrak{h} , because $T_e H = D_e = \mathfrak{h}$. Now suppose there is another subgroup \tilde{H} with the same Lie algebra. By maximality of \mathcal{H} , we have $\tilde{H} \subset H$. On the other hand, in any flat chart for D containing e , \tilde{H} must consist of

open subsets of slices. Since the slice containing e is an open subset of H , this means that \tilde{H} contains a neighborhood of e in H . But every neighborhood of e generates the connected Lie group H by Proposition 6.3.6, therefore $\tilde{H} = H$. ■

Definition 11.4.5. A **one-parameter subgroup (OPS)** of a Lie group G is a Lie group homomorphism $\gamma : (\mathbb{R}, +) \rightarrow G$, i.e. a complete curve such that $\gamma(0) = e$ and $\gamma(t+s) = \gamma(t)\gamma(s)$.

Note that a **OPS** is *not* a Lie subgroup of G , but a specific homomorphism. However the images of these homeomorphisms are weakly embedded by Theorem 11.4.2 can be identified with certain Lie subgroups:

Proposition 11.4.6. *Images of OPS's in G are precisely the connected Lie subgroups of dimension at most 1. Moreover, every such subgroup is either isomorphic to \mathbb{R} or \mathbb{S}^1 , or trivial.*

Proof. Every connected Lie subgroup of dimension 1 is automatically abelian and is the integral curve of the left-invariant vector field corresponding to any nonzero tangent vector to the subgroup at e . Such an integral curve defines the required **OPS**. The isomorphisms hold because any one-dimensional Lie group is a factor of \mathbb{R} (the only simply connected one-dimensional Lie group) by a discrete subgroup, and all such subgroups of \mathbb{R} are either trivial or isomorphic to \mathbb{Z} . ■

Theorem 11.4.7 (Characterization of one-parameter subgroups). *The OPS's of a Lie group G are precisely the maximal integral curves of left-invariant vector fields starting at $e \in G$.*

Proof. Given a left-invariant vector field, its integral curve at e will automatically be a **OPS** due to the general group property of flows and completeness (or alternatively from the uniqueness of maximal integral curves).

Conversely, every **OPS** γ is the integral curve of the left-invariant field X whose value at the identity is $X_e = \gamma'(0) \in \text{Lie } G$ because

$$\gamma'(t) = \frac{d}{ds} \Big|_{s=0} L_{\gamma(t)} \gamma(s) = L_{\gamma(t)} X_e = X_{\gamma(t)}. \quad (11.4.2)$$

This leads us to yet another common way of defining the Lie algebra of a Lie group.

Construction IV In summary, there is a natural one-to-one correspondence between elements of the Lie algebra and **OPS**'s

$$X \in \mathfrak{g} \cong \mathfrak{X}_L(G) \mapsto \gamma_X(t) = F^X(e, t). \quad (11.4.3)$$

We say that γ_X is *generated* by $X \in \mathfrak{g}$. The Lie bracket is recovered as the double derivative of the group commutator

$$[X, Y] = \frac{d}{dt} \frac{d}{ds} \Big|_{t=s=0} \gamma_X(t) \gamma_Y(s) \gamma_X(t)^{-1} \gamma_Y(s)^{-1}. \quad (11.4.4)$$

The following is a well known fact of linear algebra.

Proposition 11.4.8. *For $A \in \mathrm{gl}_n(\mathbb{K})$, the **OPS** generated by it in $\mathrm{GL}_n(\mathbb{K})$ is*

$$\gamma_A(t) = e^{tA} = \exp tA. \quad (11.4.5)$$

Corollary 11.4.9. *Uniqueness of integral curves implies that the **OPS**'s of a Lie subgroup $H \subset G$ are exactly the ones generated by $\mathfrak{h} \subset \mathfrak{g}$. Therefore in all matrix groups the **OPS**'s are given by exponentials.*

Proposition 11.4.10 ([7, Prop. 20.3]). *If $H \subset G$ is a Lie subgroup, then the **OPS**'s of H are precisely those **OPS**'s of G that are generated by elements of $\mathfrak{h} = T_e H$.*

Proof. Clearly any **OPS** $\gamma : \mathbb{R} \rightarrow H$ can also be viewed as a **OPS** of G (by composing with the inclusion map) which clearly satisfies $\gamma'(0) \in T_e H$.

Conversely, a **OPS** of G generated by an element of $T_e H$ must coincide with the analogous **OPS** of H by uniqueness of maximal integral curves. ■

11.5 Exponential map

Definition 11.5.1 (Exponential map). Given a Lie group G with Lie algebra $\mathfrak{g} = \mathrm{Lie} G$, the exponential map $\exp : \mathfrak{g} \rightarrow G$ assigns to each $X \in \mathfrak{g}$ the group element

$$\exp X = \gamma_X(1), \quad (11.5.1)$$

where γ_X is the **OPS** generated by X . As a consequence, the **OPS** generated by X can be written as

$$\gamma_X(t) = \gamma_{tX}(1) = \exp tX = e^{tX}. \quad (11.5.2)$$

Proposition 11.5.2. *The exponential map $\exp : \mathfrak{g} \rightarrow G$ has the following properties.*

- (a) \exp is smooth.
- (b) $e^{sX}e^{tX} = e^{(s+t)X}$ for $s, t \in \mathbb{R}$.
- (c) $(e^X)^{-1} = e^{-X}$.
- (d) $(e^X)^n = e^{nX}$ for $n \in \mathbb{Z}$.
- (e) $\exp_{*0} = \text{id}_{\mathfrak{g}}$.
- (f) \exp restricts to a diffeomorphism from some neighborhood of $0 \in \mathfrak{g}$ to a neighborhood of $e \in G$. Its inverse is often denoted by \ln .
- (g) \exp is natural: if H is another Lie group with $\mathfrak{h} = \text{Lie } H$, and $\Phi : G \rightarrow H$ is a Lie group homomorphism, then $\exp \circ \Phi_{*e} = \Phi \circ \exp$.
- (h) The flow F^X of a left-invariant vector field $X \in \mathfrak{X}_L(G)$ is given by $F_t = R_{e^{tx_e}}$.

Proof. (a) We need to show that the flow F^X of a vector field varies smoothly with X . This is a situation not covered by the fundamental theorem of flows, but we can reduce it to that situation by a simple trick. Define a vector field $\Xi \in \mathfrak{X}(G \times \mathfrak{g})$, where $\mathfrak{g} \cong \mathfrak{X}_L(G)$, by

$$\Xi_{(g,X)} = (X_g, 0) \in T_g G \oplus T_X \mathfrak{g}. \quad (11.5.3)$$

This is a smooth vector field because it acts on functions by differentiating them along X . It is easy to verify that the flow F^Ξ is given by

$$F_t^\Xi(g, X) = \left(F^X(g, t), X \right). \quad (11.5.4)$$

By the fundamental theorem of flows, F^Ξ is smooth. Since $\exp X = \text{pr}_1((F_1^\Xi(e, X)))$, where $\text{pr}_1 : G \times \mathfrak{g} \rightarrow G$ is the projection, it follows that \exp is smooth.

(b-d) follow trivially from the fact that $\exp(tX)$ is a OPS.

To show (e) we compute

$$\exp_{*0} X = \frac{d}{dt} \Big|_{t=0} \Phi_t^X(e) = X_e, \quad (11.5.5)$$

i.e. \exp_{*e} coincides with the natural isomorphism $\mathfrak{X}_L(G) \cong T_e G$. Then by the InMT, \exp is a local diffeomorphism and (f) follows.

(g) follows from the naturality of flows of vector fields:

$$F_t^{\Phi*X} \circ \Phi = \Phi \circ F_t^X, \quad (11.5.6)$$

and setting $t = 1$ and evaluating at the identity yields (g).

(h) was already proven in Lemma 11.1.2. ■

Remark 11.5.3. The naturality of \exp implies, in particular, that $g e^{tX} g^{-1} = e^{t \text{Ad}_g X}$. This confirms the intuitive picture that conjugate subgroups are generated by conjugate elements of the Lie algebra. For example, on SO_3 , let $H \cong \text{SO}_2$ be the OPS of rotations about the z -axis. Then if g is a rotation which maps $e_z = (0, 0, 1)$ to some vector n , then the subgroup of rotations about n is gHg^{-1} and is generated by $gL_z g^{-1}$, where L_z generates

Corollary 11.5.4. *The identity component G_0 is generated by $\exp(U)$ for any neighborhood U of the origin $0 \in \mathfrak{g}$.*

Remark 11.5.5. The inverse of the exponential map, \ln (logarithm), provides a *canonical local chart* at $e \in G$. This is a unique property of Lie groups that we will use to much benefit. We will call this chart the *logarithmic coordinates* on G .

Proposition 11.5.6. *The kernels of the adjoint representations of G and \mathfrak{g} are given by $\ker \text{Ad} = C_G(G_0)$ (centralizer of the identity component) and $\ker \text{ad} = C(\mathfrak{g})$ (center of the Lie algebra), respectively.*

Proof. The statement $\ker \text{ad} = C(\mathfrak{g})$ is obvious. Since for $g \in C_G(G_0)$ there holds $\text{Ad}_g|_{G_0} = \text{id}_{G_0}$ and since $(\text{id}_G)_{*e} = \text{id}_{\mathfrak{g}}$, we obtain $C_G(G_0) \subset \ker \text{Ad}$. Conversely, if $g \in G$ satisfies $\text{Ad}_g = \text{id}_{\mathfrak{g}}$. The naturality of the exponential map implies

$$e^X = e^{\text{Ad}_g X} = \text{Ad}_g(e^X) \text{ for all } X \in \mathfrak{g}. \quad (11.5.7)$$

That is, g commutes with all elements of $\exp(\mathfrak{g})$. Since $\exp(\mathfrak{g})$ generates G_0 , we conclude $\ker \text{Ad} \subset C_G(G_0)$. ■

Proposition 11.5.7. *If $H \subset G$ is a Lie subgroup, then the exponential map of H coincides with the restriction of the exponential map of G , and*

$$\text{Lie } H = \{X \in \text{Lie } G : \exp(tX) \in H \text{ for all } t \in \mathbb{R}\}. \quad (11.5.8)$$

Proof. The fact that $\exp_H = \exp_G|_{\mathfrak{h}}$ follows immediately from Proposition 11.4.10.

Next assume that $e^{tX} \in H$ for all t . Since H is weakly embedded in G by Theorem 11.4.2, it follows that the curve $t \mapsto e^{tX}$ is smooth as a map into H , and thus $X = \gamma'(0) \in T_e H$. Conversely, if $X \in T_e H$ then Proposition 11.4.10 implies that $e^{tX} \in H$ for all t . ■

Theorem 11.5.8. *The following formulas hold for all $X \in \mathfrak{g}, g \in G$:*

- (a) $\text{Ad}_e X = e^{\text{ad}_X}$.
- (b) $\text{Ad}_g e^X = e^{\text{Ad}_g X}$.

Proof. For (a) we apply the naturality of \exp (Proposition 11.5.2 (g)) to the homomorphism $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$, and for (b) we apply it to the homomorphism $\text{Ad}_g : G \rightarrow G$. ■

Proposition 11.5.9. *If $X, Y \in \mathfrak{g}$ commute, i.e. $[X, Y] = 0$, then*

$$e^X e^Y = e^Y e^X = e^{X+Y}. \quad (11.5.9)$$

In particular, if \mathfrak{g} is abelian, \exp is a Lie group homomorphism from the vector group underlying \mathfrak{g} to G .

Proof. It suffices to show that the curve defined by $\gamma(t) = e^{tX} e^{tY}$ coincides with the integral curve generated by $X + Y$. Indeed, $\gamma(0) = e$ and we can use the commutation of the flows of the left-invariant fields X_L and Y_L to compute the generator. Here we will need the general product rule that holds for maps $f : M_1 \times M_2 \rightarrow N$:

$$T_{(m_1, m_2)} f(X_1, X_2) = T_{m_1} f_{m_2}(X_1) + T_{m_2} f_{m_1}(X_2) \quad (11.5.10)$$

where $f_{m_1} : M_2 \rightarrow N$ and $f_{m_2} : M_1 \rightarrow N$ are the partial mappings. In particular, if $M_1 = M_2 \cong \mathbb{R}$ then we have the “product rule” for derivatives along the diagonal:

$$\frac{d}{dt} \Big|_t \varphi(t, t) = \frac{d}{ds} \Big|_{s=t} \varphi(s, t) + \frac{d}{ds} \Big|_{s=t} \varphi(t, s). \quad (11.5.11)$$

Applying this to γ we get

$$\gamma'(0) = \frac{d}{dt} \Big|_{t=0} F_t^{Y_L} \circ F_t^{X_L}(e) = \frac{d}{ds} \Big|_{s=0} F_s^{Y_L}(e) + \frac{d}{ds} \Big|_{s=0} F_s^{X_L}(e) = Y_e + X_e. \quad (11.5.12)$$

When the elements do not commute, the following two propositions provide an approximate formula. ■

Proposition 11.5.10 ([7, Prop. 20.10]). *For any $X, Y \in \mathfrak{g} = \text{Lie } G$, there is a smooth function $Z : (-\epsilon, \epsilon) \rightarrow \mathfrak{g}$ for some $\epsilon > 0$ such that*

$$e^{tX} e^{tY} = \exp \left(t(X + Y) + t^2 Z(t) \right) \quad (11.5.13)$$

for all $t \in (-\epsilon, \epsilon)$.

Proof. Since \exp is a local diffeomorphism, there exists an $\epsilon > 0$ such that on $(-\epsilon, \epsilon)$ we can consider the smooth map

$$\varphi(t) = \ln(\mathrm{e}^{tX} \mathrm{e}^{tY}). \quad (11.5.14)$$

Obviously $\varphi(0) = 0$ and since the differential of the multiplication map is $T_{(e,e)}m(X, Y) = X + Y$, we have

$$\varphi'(0) = \ln_{*e}(X + Y) = X + Y. \quad (11.5.15)$$

By Taylor's theorem $\varphi(t) = t(X + Y) + t^2 Z_2(t)$ for some smooth $Z_3(t)$. ■

Corollary 11.5.11. *Under the hypotheses of the preceding Proposition,*

$$\lim_{n \rightarrow \infty} \left(\exp \frac{tX}{n} \exp \frac{tY}{n} \right)^n = \exp t(X + Y). \quad (11.5.16)$$

Proof. We have

$$\begin{aligned} \left(\exp \frac{tX}{n} \exp \frac{tY}{n} \right)^n &= \left(\exp \left(\frac{t}{n}(X + Y) + \frac{t^2}{n^2} Z_2\left(\frac{t}{n}\right) \right) \right)^n = \\ &= \exp \left(t(X + Y) + \frac{t^2}{n^2} Z_2\left(\frac{t}{n}\right) \right). \end{aligned} \quad (11.5.17)$$

Taking the limit $n \rightarrow \infty$ at fixed t gives the result. ■

Proposition 11.5.12. *Under the hypotheses of the preceding Proposition,*

$$\mathrm{e}^{tX} \mathrm{e}^{tY} = \exp \left(t(X + Y) + \frac{t^2}{2} [X, Y] + t^3 Z_3(t) \right) \quad (11.5.18)$$

for another smooth map $Z_3 : (-\epsilon, \epsilon) \rightarrow \mathfrak{g}$.

Proof. For the function φ from the proof of Proposition 11.5.10 we have

$$\varphi(t) = t(X + Y) + t^2 Z_2(0) + t^3 Z_3(t). \quad (11.5.19)$$

To compute the flow of this vector field we can apply the general Taylor formula for manifolds (9.4.12). By an abuse of notation we drop the smooth function $f \in C^\infty(G)$ on which this formula is supposed to act:

$$\mathrm{e}^{\varphi(t)} = F_t^{X+Y+\frac{s}{2}Z_2(0)+O(s^2)} \Big|_{s=t} = 1 + t(X + Y) + \frac{t^2}{2} \left((X + Y)^2 + 2Z_2(0) \right) + O(t^3). \quad (11.5.20)$$

On the other hand the Taylor formula for two fields (9.4.13) gives

$$e^{tX} e^{tY} = 1 + t(X + Y) + \frac{t^2}{2}(X^2 + 2XY + Y^2) + O(t^3). \quad (11.5.21)$$

We read off $Z_2 = [X, Y]/2$. ■

This gives us a new interpretation of the Lie bracket as the commutator of infinitesimal flows of left-invariant vector fields (closely connected to formula (6.10.9)).

Corollary 11.5.13. *By a repeated application of the Taylor series obtained above,*

$$e^{tX} e^{tY} e^{-tX} = e^{tY + t^2[X, Y] + O(t^3)} \quad (11.5.22)$$

and

$$[e^{tX}, e^{tY}] = e^{tX} e^{tY} e^{-tX} e^{-tY} = e^{t^2[X, Y] + O(t^3)}. \quad (11.5.23)$$

Remark 11.5.14. By continuing the Taylor expansion endlessly, we can obtain the full formal series for $\ln(e^X e^Y)$ known as the *Baker-Campbell-Hausdorff formula*. Its special property is that it can be written using only Lie brackets (or adjoints). We will not derive the full series here.

11.6 Differential of exp

Theorem 11.6.1 ([33, Thm. 1.5.3]). *For any $X \in \mathfrak{g}$ the linear mapping $\exp_{*X} : \mathfrak{g} \rightarrow T_{e^X} G$ is given by*

$$\exp_{*X} = (R_{e^X})_{*e} \circ \int_0^1 e^{s\text{ad}_X} ds = (L_{e^X})_{*e} \circ \int_0^1 e^{-s\text{ad}_X} ds. \quad (11.6.1)$$

Proof. We will provide a simple derivation of this from the Maurer-Cartan equation (11.1.12). Let us apply it to the map

$$g(t, s) = e^{t(X+sY)}. \quad (11.6.2)$$

Then we find

$$\xi = \left(R_g^{-1} \right)_* \partial_t g = X + sY, \quad \eta = \left(R_g^{-1} \right)_* \partial_s g, \quad (11.6.3)$$

and the Maurer-Cartan equation $\partial_t \eta - \partial_s \xi = [\xi, \eta]$ now reads

$$\partial_t \eta = Y + [X, \eta]. \quad (11.6.4)$$

By denoting $\gamma(t) = \eta(t, 0)$, we get the differential equation at $s = 0$

$$\dot{\gamma}(t) = Y + [X, \gamma]. \quad (11.6.5)$$

First we notice that the solution of the homogenous equation $\dot{\gamma} = [X, \gamma]$ is given by the adjoint curves

$$\gamma_{\text{hom}} = \text{Ad}_{e^{tX}} \gamma_0 = e^{t\text{ad}_X} \gamma_0. \quad (11.6.6)$$

By Lagrange's method of variation of constants, we look for a solution to (11.6.5) in the form $\gamma(t) = e^{t\text{ad}_X} \zeta(t)$. Plugging this ansatz back into the equation gives

$$e^{t\text{ad}_X} \dot{\zeta}(t) = Y. \quad (11.6.7)$$

The solution of this with $\gamma(0) = 0$ is obviously given by

$$\zeta(t) = \int_0^t e^{-s\text{ad}_X} Y ds. \quad (11.6.8)$$

On the other hand, from the definition of γ ,

$$\gamma(t) = \left(R_{e^{tX}}^{-1} \right)_{*e^{tX}} \circ \exp_{*e^{tX}}(tY) \quad (11.6.9)$$

and thus

$$\exp_{*e^{tX}}(tY) = (L_{e^{tX}})_{*e} \int_0^t e^{-s\text{ad}_X} ds (tY). \quad (11.6.10)$$

By setting $t = 1$ we get the claimed formula. The second version is obtained via the substitution $s = 1 - s'$ in the integral. ■

Remark 11.6.2 (Singular points of \exp). Since the integral derived above is elementary, the differential of \exp at $X \in \mathfrak{g}$ can be given by the formal series

$$(L_{e^{-X}})_{*e} \circ \exp_{*X} = \sum_{n=0}^{\infty} \frac{(-\text{ad}_X)^n}{(n+1)!} = \frac{I - e^{-\text{ad}_X}}{\text{ad}_X}, \quad (11.6.11)$$

$$(R_{e^{-X}})_{*e} \circ \exp_{*X} = \sum_{n=0}^{\infty} \frac{(\text{ad}_X)^n}{(n+1)!} = \frac{e^{\text{ad}_X} - I}{\text{ad}_X}, \quad (11.6.12)$$

where in case of non-invertible matrices ad_X the series should be taken as the *definition* of this latter expression. Indeed ad_X is never invertible because it always has X itself in its kernel, $\text{ad}_X X = 0$. By looking at the location of the zeros of the function $f(z) = (e^z - 1)/z$, we conclude that the differential \exp_{*X} is bijective iff no eigenvalue of ad_X lies in $2\pi i\mathbb{Z} \setminus \{0\}$. Conversely, the singular points of the exponential map are precisely those $X \in \mathfrak{g}$ such that $\text{ad}_X \in \text{End}(V)$ has an eigenvalue of the form $2\pi i k$, $k \in \mathbb{Z} \setminus \{0\}$.

Example 11.6.3 (Exponential map of $\text{SO}_3(\mathbb{R})$). Recall that the rotation group $G = \text{SO}_3(\mathbb{R})$ can be parametrized by vectors $n \in \mathbb{R}^3$ as follows:

$$R_n = I + \frac{\sin |n|}{|n|} A_n + \frac{1 - \cos |n|}{|n|} A_n^2, \quad (11.6.13)$$

where $|n| = \sqrt{n_1^2 + n_2^2 + n_3^2}$ is the rotation angle and

$$A_\omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \quad (11.6.14)$$

The conjugacy classes of $\text{SO}_3(\mathbb{R})$ are exactly the spheres $C_{|n|}$ of equal rotation angle $0 \leq |n| \leq \pi$ (except $C_0 = \{I\}$). Each conjugacy class in the range $0 < |n| < \pi$ is diffeomorphic to a sphere, whereas C_π , the conjugacy class of the matrix $\text{diag}(1, -1, -1)$, is diffeomorphic to \mathbb{RP}^2 .

Then the singular points of the exponential map are the antisymmetric matrices $A_\omega \in \mathfrak{so}_3(\mathbb{R})$ with $|\omega| = 2\pi$. In this case $e^{A_\omega} = I \in \text{SO}_3(\mathbb{R})$.

The exponential map is a local diffeomorphism from the ball $|\omega| < 2\pi$ onto $\text{SO}_3(\mathbb{R})$ which restricts to a two-fold covering from the punctured ball $0 < |\omega| < 2\pi$ onto $\text{SO}_3(\mathbb{R}) \setminus \{I\}$. The ball $|\omega| < \pi$ is the maximal open subset of $\mathfrak{so}_3(\mathbb{R})$ on which the exponential mapping is injective; it is mapped diffeomorphically onto the complement in $\text{SO}_3(\mathbb{R})$ of the conjugacy class C_π .

Example 11.6.4 (Exponential map of $\text{SL}_2(\mathbb{C})$ and $\text{SU}_2(\mathbb{C})$). If $G = \text{SL}_2(\mathbb{C})$, then its Lie algebra $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ consists of traceless matrices;

$$X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \quad a, b, c \in \mathbb{C}, \quad (11.6.15)$$

and any such matrix satisfies the equation

$$X^2 = \lambda^2 I, \quad \lambda = \pm \sqrt{a^2 + bc}, \quad (11.6.16)$$

where λ are the eigenvalues of X . Due to this the exponential series separates into two sums over odd and even powers, and we get

$$e^X = \cosh \lambda \cdot I + \frac{\sinh \lambda}{\lambda} X. \quad (11.6.17)$$

Observe that the coefficients here are entire analytic functions of $\lambda^2 = a^2 + bc$ and hence are

entire analytic functions on $\mathfrak{sl}_2(\mathbb{C})$. If $\lambda = 0$, we recover the result

$$e^X = 1 + X \Leftrightarrow X \text{ is nilpotent} \quad (11.6.18)$$

(because for 2×2 matrices nilpotency is equivalent to the vanishing of the square). If $\lambda = it$ with $t \in \mathbb{R}$, the exponential becomes

$$e^X = \cos t \cdot I + \frac{\sin t}{t} X. \quad (11.6.19)$$

This occurs if $a, b, c \in \mathbb{R}$ (which is equivalent to $X \in \mathfrak{sl}_2(\mathbb{R})$) and $a^2 + bc < 0$, i.e. “ X is in the elliptic domain”. The case $\lambda = it$, $t \in \mathbb{R}$ also occurs for all matrices of the form

$$X = \begin{pmatrix} i\alpha & \beta \\ -\bar{\beta} & -i\alpha \end{pmatrix} \in \mathfrak{su}_2. \quad (11.6.20)$$

In this case $t = \sqrt{a^2 + |\beta|^2}$, which is the Euclidean length of $(\alpha, \beta) \in \mathbb{R} \times \mathbb{C} \cong \mathbb{R}^3$. For $X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ we recover the usual notation

$$R(t) = \exp \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R}). \quad (11.6.21)$$

If $X \in \mathfrak{sl}_2(\mathbb{R})$, then e^X has eigenvalues equal to -1 iff there is a $k \in \mathbb{Z}$ such that X has eigenvalues equal to $\pm(i\pi + 2\pi ik)$ (with both signs occurring). This implies that X is diagonalizable over \mathbb{C} , hence e^X is as well, so $e^X = -I$. It follows that elements $A \in \mathrm{SL}_2(\mathbb{R})$ that have eigenvalues equal to -1 without themselves being equal to $-I$, that is, those $A \in \mathrm{SL}_2(\mathbb{R})$ which are conjugate to $\begin{pmatrix} -1 & \pm 1 \\ 0 & -1 \end{pmatrix}$, are not in the image of $\exp : \mathfrak{sl}_2(\mathbb{R}) \rightarrow \mathrm{SL}_2(\mathbb{R})$ (and only they aren't).

Defining the algebraic variety in \mathfrak{g}

$$\Sigma_1 = \{X \in \mathfrak{g} \mid \det((\mathrm{ad}_X)_{\mathbb{C}} - 2\pi i I)\}, \quad (11.6.22)$$

the set of all singular points of the exponential map is $\Sigma = \bigcup_{k \in \mathbb{C} \setminus \{0\}} k\Sigma_1$. In the case of $\mathrm{SL}_2(\mathbb{C})$,

$$\Sigma_1 = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{C}) \mid a^2 + bc = -\pi^2 \right\}. \quad (11.6.23)$$

Thus for $G = \mathrm{SU}_2$,

$$\Sigma_1 = \left\{ \begin{pmatrix} i\alpha & \beta \\ -\bar{\beta} & -i\alpha \end{pmatrix} \in \mathfrak{su}_2 \mid a^2 + |\beta|^2 = \pi^2 \right\}, \quad (11.6.24)$$

which is a sphere \mathbb{S}^2 in \mathfrak{su}_2 of Euclidean radius π , and \exp maps the ball $a^2 + |\beta|^2 < \pi^2$ diffeomorphically onto $\mathrm{SU}_2 \setminus \{-I\}$. On the other hand, for $G = \mathrm{SL}_2(\mathbb{R})$, the set Σ_1 is the two-sheeted hyperboloid

$$\Sigma_1 = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{R}) \mid a^2 + bc = -\pi^2 \right\}, \quad (11.6.25)$$

mapped to $\{-I\}$ by the exponential. Here \exp is a diffeomorphism from the set $a^2 + bc > -\pi^2$ onto the complement in $\mathrm{SL}_2(\mathbb{R})$ of $\{-I\}$ and the conjugacy classes of $\begin{pmatrix} -1 & \pm 1 \\ 0 & -1 \end{pmatrix}$, the latter being the elements that are not in the image of $\exp : \mathfrak{sl}_2(\mathbb{R}) \rightarrow \mathrm{SL}_2(\mathbb{R})$ at all. Note that $\exp_{\mathfrak{sl}_2}(\mathbb{R})$ is neither open in $\mathrm{SL}_2(\mathbb{R})$ ($-I$ belongs to it, but not as an interior point), nor closed.

Exercise 11.4. Show that the exponential map $\exp : \mathfrak{gl}_n(\mathbb{C}) \rightarrow \mathrm{GL}_n(\mathbb{C})$ is surjective. Hint: apply the Jordan canonical form to write a conjugate of $A \in \mathrm{GL}_n(\mathbb{C})$ as $D(I + U)$ with $D \in \mathrm{GL}_n(\mathbb{C})$ diagonal and U nilpotent, and use $\ln(I + U) = \sum_{k \geq 0} \frac{(-1)^k}{k+1} U^{k+1}$, which is a terminating series.

11.7 Normal subgroups

Normal subgroups (those that are invariant under conjugation) play a central role in abstract group theory: they are the only subgroups whose quotients have group structures, and the only subgroups that are kernels of group homomorphisms. First we derive a useful criterion for normality.

Lemma 11.7.1 ([7, Lem. 20.23]). *Let G be a connected Lie group and $H < G$ a connected Lie subgroup. Let \mathfrak{g} and \mathfrak{h} denote their respective Lie algebras. Then H is normal in G iff*

$$e^X e^Y e^{-X} \text{ for all } X \in \mathfrak{g}, Y \in \mathfrak{h}. \quad (11.7.1)$$

Proof. Since $(e^X)^{-1} = e^{-X}$, if H is normal, the property in question holds automatically. Conversely, choose open neighborhoods $V \subset \mathfrak{g}$ of the origin and $U \subset G$ of the identity such that $\exp : V \rightarrow U$ is a diffeomorphism. Since the exponential map of H is the restriction of that of G , after shrinking V if necessary, we may assume that the restriction of \exp to $V \cap \mathfrak{h}$ is a diffeomorphism from $V \cap \mathfrak{h}$ to a neighborhood U_0 of the identity in H . Shrinking V still further, we may assume that $X \in V$ iff $-X \in V$. Then the assumption implies that $ghg^{-1} \in H$ whenever $g \in U$ and $h \in U_0$.

Since every element of H can be written as a finite product $h = h_1 \cdots h_m$ with $h_1, \dots, h_m \in U_0$ (Proposition 6.3.6), it follows that for any $g \in U$ and $h \in H$ we have

$$ghg^{-1} = gh_1 \cdots h_m g^{-1} = (gh_1g^{-1}) \cdots (gh_mg^{-1}) \in H. \quad (11.7.2)$$

Similarly, any $g \in G$ can be written $g = g_1 \cdots g_k$ with $g_1, \dots, g_k \in U$, so it follows by induction on k that $ghg^{-1} \in H$ for all $g \in G, h \in H$. ■

We can now establish a relationship between normal subgroups and ideals of the Lie algebra.

Theorem 11.7.2 (Ideals and Normal Subgroups[7, Thm. 20.28]). *Let G be a connected Lie group, and suppose $H < G$ is a connected Lie subgroup. Then H is a normal subgroup of G iff $\text{Lie } H$ is an ideal in $\text{Lie } G$.*

Proof. Write \mathfrak{g} and \mathfrak{h} for the two Lie algebras, treating \mathfrak{h} as a subspace of \mathfrak{g} . For any $g \in G$, we have the equality

$$\mathbf{Ad}_g \circ \exp = \exp \circ \mathbf{Ad}_g. \quad (11.7.3)$$

Suppose that \mathfrak{h} is an ideal. Applying this equation to $Y \in \mathfrak{h}$ with $g = \exp X$, we obtain

$$\exp(\mathbf{Ad}_{e^X} Y) = \mathbf{Ad}_{e^X}(e^Y) = e^X e^Y e^{-Y}. \quad (11.7.4)$$

On the other hand, we also have the equation

$$\mathbf{Ad}_{e^X} = e^{\mathbf{ad}_X} = \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{ad}_X)^k. \quad (11.7.5)$$

Whenever $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, we have $\mathbf{ad}_X Y = [X, Y] \in \mathfrak{h}$, and by induction $(\mathbf{ad}_X)^k Y \in \mathfrak{h}$ for all k . Since finite-dimensional vector spaces are closed, the series converges to an element of \mathfrak{h} and thus $\mathbf{Ad}_{e^X} Y \in \mathfrak{h}$, and (11.7.4) implies that $e^X e^Y e^{-X} \in \exp(\mathfrak{h}) \subset H$. By Lemma 11.7.1, H is normal.

Conversely, suppose H is normal. Given $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, note that (11.7.3) applied to sY with $g = e^{tX}$ implies

$$\exp(\mathbf{Ad}_{e^{tX}} sY) = e^{tX} e^{sY} e^{-tX} \in H. \quad (11.7.6)$$

Since $\mathbf{Ad}_{e^{tX}}$ is linear, it follows that

$$\exp(s \mathbf{Ad}_{e^{tX}} Y) = \exp(\mathbf{Ad}_{e^{tX}} sY), \quad (11.7.7)$$

which we have just shown to be in H for all s , so $\mathbf{Ad}_{e^{tX}} Y \in \mathfrak{h}$ by Proposition 11.5.7. Since

$$\left. \frac{d}{dt} \right|_{t=0} \mathbf{Ad}_{e^{tX}} Y = \mathbf{ad}_X Y = [X, Y], \quad (11.7.8)$$

we conclude that $[X, Y] \in \mathfrak{h}$ and \mathfrak{h} is an ideal. ■

11.8 Covering groups

Lie group isomorphisms induce isomorphisms of Lie algebras. Now we ask to which extent a Lie group is determined by its Lie algebra.

Theorem 11.8.1 ([40, Thm. 9.5.13]). *Two connected Lie groups have isomorphic Lie algebras iff their universal covering groups are isomorphic.*

Proof. The forward implication is obvious because the universal covering group has the same Lie algebra as the original group. Conversely, let $\phi : \mathfrak{h} \rightarrow \mathfrak{h}$ be an isomorphism. Using the Integrability Theorem we obtain a unique homomorphism $\Phi : \tilde{G} \rightarrow \tilde{H}$ with $\Phi_{*e} = \phi$ and also a unique morphism $\Psi : \tilde{H} \rightarrow \tilde{G}$ with $\Psi_{*e} = \phi^{-1}$. Then $(\Phi \circ \Psi) = \text{id}_{\mathfrak{g}}$ implies $\Phi \circ \Psi = \text{id}_{\tilde{H}}$ and likewise $\Psi \circ \Phi = \text{id}_{\tilde{G}}$. Therefore $\tilde{G} \cong \tilde{H}$. ■

Combining this with the fact that every Lie group is the quotient of its universal covering group by a discrete normal (and central) subgroup $\Gamma \subset \tilde{G}$ (Corollary 6.2.9), we obtain:

Corollary 11.8.2. *Let G be a connected Lie group with universal covering homomorphism $\pi : \tilde{G} \rightarrow G$. Then for each discrete central subgroup $\Gamma \subset \tilde{G}$, the group \tilde{G}/Γ is a connected Lie group with $\text{Lie}(\tilde{G}/\Gamma) \cong \mathfrak{g}$ and, conversely, each Lie group with the same Lie algebra as G is isomorphic to some quotient \tilde{G}/Γ , so that $\Gamma \cong \pi_1(G)$.*

Example 11.8.3. It's important which specific subgroup $\Gamma \subset \tilde{G}$ one chooses, not just its isomorphism class. We describe a pair of non-isomorphic Lie groups with isomorphic Lie algebras and isomorphic fundamental groups.

Let $\tilde{G} = \text{SU}_2 \times \text{SU}_2$ whose center is $\mathbb{Z}_2 \times \mathbb{Z}_2$, and

$$G = \tilde{G}/(\mathbb{Z}_2 \times \{I\}) \cong \text{SO}_3 \times \text{SU}_2, \quad H = \tilde{G}/\{\pm(I, I)\} \cong \text{SO}_4, \quad (11.8.1)$$

where the latter isomorphism was demonstrated in Example 6.4.3. Then $\pi_1(G) \cong \pi_1(H) \cong \mathbb{Z}_2$, but there is no automorphism of \tilde{G} mapping $\pi_1(G)$ to $\pi_1(H)$.

Indeed, one can show that the two direct factors are the only nontrivial connected normal subgroups of \tilde{G} , so that each automorphism of \tilde{G} either preserves both or exchanges them. Since $\pi_1(H)$ is not contained in any of them, it cannot be mapped to $\pi_1(G)$ by an automorphism of \tilde{G} .

Example 11.8.4. Here are some examples of pairs of linear Lie groups with isomorphic Lie algebras:

1. SO_3 and SU_2 , see Example 6.1.10.
2. $(\text{SO}_{2,1})_0$ and $\text{SL}_2(\mathbb{R})$. Just like in the previous case, there is a covering homomorphism $\Phi : H \rightarrow G$ coming from the adjoint representation $\text{Ad} : \text{SL}_2(\mathbb{R}) \rightarrow \text{GL}(\text{Lie } \text{SL}_2(\mathbb{R})) \cong$

$\mathrm{GL}_3(\mathbb{R})$.

3. $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SO}_{3,1}(\mathbb{R})_0$, see Example 6.4.5.

4. Lie groups that are homeomorphic as spaces don't have to be isomorphic as groups. As an example, SO_4 is homeomorphic to $\mathrm{SO}_3 \times \mathrm{Sp}_1$, whereas the only related isomorphism is $\mathrm{SO}_4 \cong \mathrm{Sp}_1 \times \mathrm{Sp}_1 / \{\pm I\}$. Another example is the homeomorphic pair U_2 and $\mathrm{SU}_2 \times U_1$. However, if two Lie groups are homeomorphic and their Lie algebras are isomorphic, then the groups are isomorphic.

Determining which of the groups \tilde{G}/Γ in general is a very subtle question. We now demonstrate an example of a nonlinear Lie group.

Recall that a Banach algebra is a complete normed vector space with a multiplication operation such that $\|xy\| \leq \|x\|\|y\|$. Any finite-dimensional algebra can be represented as a sub-algebra of some matrix algebra $\mathrm{Mat}_n(\mathbb{K})$, and the commutator of any two matrices is traceless. The following analog of this statement on general Banach algebras turns out to hold.

Lemma 11.8.5. *If A is a Banach algebra with unit $\mathbf{1}$ and $p, q \in A$ with $[p, q] = pq - qp = \lambda \mathbf{1}$, then $\lambda = 0$.*

Proof. By induction we obtain $[p, q^n] = \lambda n q^{n-1}$ for $n \in \mathbb{N}$. Therefore

$$n|\lambda| \|q^{n-1}\| \leq 2\|p\| \|q^n\| \leq 2\|p\| \|q\| \|q^{n-1}\|, \quad (11.8.2)$$

which leads to

$$(n|\lambda| - 2\|p\| \|q\|) \|q^{n-1}\| \leq 0. \quad (11.8.3)$$

If $\lambda \neq 0$, then we obtain for sufficiently large n that $q^{n-1} = 0$. Thus for $n > 1$, we derive that $q^{n-2} = 0$. By induction, $q = 0$ and hence $\lambda = 0$. ■

Example 11.8.6 (Quantum Heisenberg group). The Heisenberg group from Example 6.11.5 contains a normal subgroup consisting of matrices of the form

$$\begin{pmatrix} 1 & 0 & n\hbar \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad n \in \mathbb{Z}, \quad (11.8.4)$$

where $\hbar > 0$ is a fixed constant. In fact this is a discrete central subgroup $D \subset H_3$, isomorphic to \mathbb{Z} , which gives the exact sequence

$$0 \rightarrow \mathbb{R}/\hbar\mathbb{Z} \hookrightarrow H_3/\hbar\mathbb{Z} \xrightarrow{p} \mathbb{R}^2 \rightarrow 0, \quad (11.8.5)$$

where of course \mathbb{R}/\mathbb{Z} can be identified with the circle group U_1 and p is the homomorphism

$$p : \begin{pmatrix} 1 & x & z + \hbar\mathbb{Z} \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mapsto (x, y) \in \mathbb{R}^2 \quad (11.8.6)$$

The group $H_3/\hbar\mathbb{Z}$ is also often called the Heisenberg group (it appears in the context of quantum mechanics) and it is nonlinear, as we will show now.

Let $G = H_3$. Note that \exp_G is a diffeomorphism whose inverse is given by

$$\ln(g) = (g - I) - \frac{1}{2}(g - I)^2, \quad (11.8.7)$$

see Corollary 12.2.2 below. Recall the commutation relations $[X, Y] = Z$, $[Y, Z] = [X, Z] = 0$ (Example 6.11.5). Therefore, by the BCH formula,

$$\exp(\mathbb{R}Z) = I + \mathbb{R}Z \subset Z(G), \quad (11.8.8)$$

and $D = \exp(\hbar\mathbb{Z}Z)$ is a discrete central subgroup of G . We claim that G/D is nonlinear. Consider a homomorphism $\Phi : G \rightarrow \mathrm{GL}_n(\mathbb{C})$ with $D \subset \ker \Phi$ (so it induces a finite-dimensional representation of G/D). Let $\phi = \Phi_{*e} : \mathfrak{g} \rightarrow \mathfrak{gl}_n(\mathbb{C})$, so we obtain matrices

$$P := \phi(X), \quad Q := \phi(Y), \quad C := \phi(\hbar Z) \quad (11.8.9)$$

with $[P, Q] = C/\hbar$. Now $e^{\hbar Z} \in D \subset \ker \Phi$ implies that $e^C = \Phi(e^{\hbar Z}) = I$, and hence that C is diagonalizable with all eigenvalues contained in $2\pi i\mathbb{Z}$. Let $V_\lambda = \ker(C - \lambda I)$. Since Z is central in \mathfrak{g} , the space V_λ is invariant under G , hence also under \mathfrak{g} . Therefore the restrictions $P_\lambda = P|_{V_\lambda}$ and $Q_\lambda = Q|_{V_\lambda}$ satisfy $[P_\lambda, Q_\lambda] = \lambda \mathrm{id}$ in the Banach algebra $\mathrm{End}(V_\lambda)$. In view of the preceding lemma, we have $\lambda = 0$. Therefore, the diagonalizability of C entails that $C = 0$ and hence that $\mathbb{R}Z \subset \ker \phi$. It follows in particular that the group G/D has no faithful finite-dimensional linear representations.

11.9 (*) Technical lemmas

Here we provide the more conventional, but longer, proof of Theorem 11.6.1 that doesn't use the Maurer-Cartan equation. It is essentially the same proof, but it rederives the basic results on developments of Lie algebra-valued 1-forms from Section 11.3 as statements about the existence of solutions for certain ODEs. First we will need the following technical lemma about ODEs dependent on a parameter.

Lemma 11.9.1 ([33, Lem. B.4, Thm. B.3]). Consider an ODE of the form $\dot{\gamma}(t) = f(\gamma(t))$ on some domain in \mathbb{R}^n with a differentiable f . Fix an initial value $\gamma(0) = x$. As is well known, this problem has a unique solution $\gamma(t, x)$ defined for sufficiently small t , C^1 in t , and of the same smoothness class as f in x . If f in addition depends differentiably on a parameter ϵ , then

$$\partial_\epsilon \gamma(t, x, \epsilon) = \int_0^t \frac{\partial \gamma}{\partial x}(t-s, \gamma(s, x, \epsilon), \epsilon) \frac{\partial f}{\partial \epsilon}(\gamma(s, x, \epsilon), \epsilon) ds. \quad (11.9.1)$$

On a manifold M , if $X_\epsilon \in \mathfrak{X}(M)$ is a vector field that smoothly depends on a parameter ϵ , then its flow F_t^ϵ also depends smoothly on ϵ and satisfies

$$\partial_\epsilon F_t^\epsilon(m) = \int_0^t (T_{F_s^\epsilon(m)} F_{t-s}^\epsilon) \cdot \frac{\partial X_\epsilon}{\partial \epsilon}(F_s^\epsilon(m)) ds \in T_{F_t^\epsilon(m)} M. \quad (11.9.2)$$

Proof. First we rewrite the ODE in the form of an integral equation

$$\gamma(t) = x + \int_0^t f(\gamma(s)) ds. \quad (11.9.3)$$

Differentiating this with respect to ϵ we get

$$\partial_\epsilon \gamma(t, x, \epsilon) = \int_0^t (\partial_x f(\gamma(s, x, \epsilon), \epsilon) \partial_\epsilon \gamma(s, x, \epsilon) + \partial_\epsilon f(\gamma(s, x, \epsilon), \epsilon)) ds. \quad (11.9.4)$$

Differentiating this with respect to t we see that $\partial_\epsilon \gamma$ satisfies the inhomogenous linear differential equation

$$\partial_\epsilon \dot{\gamma}(t) = \partial_x f(\gamma(t, x, \epsilon), \epsilon) \partial_\epsilon \gamma(t, x, \epsilon) + \partial_\epsilon f(\gamma(t, x, \epsilon), \epsilon), \quad \partial_\epsilon \gamma(0, x, \epsilon) = 0. \quad (11.9.5)$$

Meanwhile differentiating the original ODE $\dot{\gamma}(t) = f(\gamma(t))$ with respect to x gives

$$\partial_x \dot{\gamma}(t, x, \epsilon) = \partial_x f(\gamma(t, x, \epsilon), \epsilon) \partial_x \gamma(t, x, \epsilon), \quad (11.9.6)$$

and therefore $t \mapsto \partial_x \gamma(t, x, \epsilon)$ is a solution of the homogenous equation associated with (11.9.5). The classical variation-of-constants formula of Lagrange yields the claimed formula as the solution of (11.9.5).

The formula on manifolds is just the same formula if $|t|$ is sufficiently small. It extends to the full flow domain of X_ϵ by “continuous induction” in t , namely applying the chain rule to the group law of flows, $F_{t+\tau}^\epsilon = F_\tau^\epsilon \circ F_t^\epsilon$. One assumes that the formula already holds for $|t| < \delta$ and then replaces x with $F_t^\epsilon(x)$ and t with τ , $|\tau| < \delta$. Unpacking $\partial_\epsilon F_{t+\tau}^\epsilon$ via a double application of the formula then confirms that the formula extends to $t + \tau$. ■

Second proof of Theorem 11.6.1. Now we apply formula (11.9.2) to $s \mapsto R_{e^{s(X+\epsilon Y)}}$, which is the flow F_s^ϵ of the ϵ -dependent vector field $(X + \epsilon Y)_L$ on G . Hence:

$$\begin{aligned} \exp_{*X} Y &= \frac{d}{d\epsilon} \Big|_{\epsilon=0} \exp(X + \epsilon Y) = \frac{d}{d\epsilon} \Big|_{\epsilon=0} R_{\exp(X+\epsilon Y)}(e) = \\ &= \int_0^1 T_{e^{sX}} R_{e^{(1-s)X}} \cdot Y_L(e^{sX}) ds = \int_0^1 T_e R_{e^{sX}} \circ T_{e^{sX}} R_{e^{-sX}} \circ T_e L_{e^{sX}} \cdot Y ds = \\ &= T_e R_{e^X} \circ \int_0^1 T_e (R_{e^{-sX}} \circ L_{e^{sX}}) ds \cdot Y = T_e R_{e^X} \circ \int_0^1 \text{Ad}_{e^{sX}} ds \cdot Y = \\ &= T_e R_{e^X} \circ \int_0^1 e^{s\text{ad}_X} ds \cdot Y. \quad (11.9.7) \end{aligned}$$

In the last equality we used Theorem 11.5.8. This proves the first identity. The second one follows from:

$$\begin{aligned} T_e R_{e^X}^{-1} \circ T_e L_{e^X} \circ \int_0^1 e^{-s\text{ad}_X} ds &= \\ &= \text{Ad}_{e^X} \circ \int_0^1 e^{-s\text{ad}_X} ds = \int_0^1 e^{(1-s)\text{ad}_X} ds = \int_0^1 e^{u\text{ad}_X} du, \quad (11.9.8) \end{aligned}$$

by means of the substitution $1 - s = u$. ■

The next technical lemma is useful in many important theorems, and we've already used it in the proof of Theorem 11.3.10.

Definition 11.9.2 (Path space). The space of all paths in a smooth manifold M starting at $x_0 \in M$ is denoted by $P(M, x_0)$ and is provided with the topology of uniform convergence on M . Paths are considered equivalent if they are homotopic with fixed ends.

The path space $P(\mathfrak{g})$ of a Lie algebra \mathfrak{g} (or any finite-dimensional vector space) is the space of all continuous paths $\delta : [0, 1] \rightarrow \mathfrak{g}$ with the topology of uniform convergence defined by the supremum under some norm on \mathfrak{g} (all choices of norms are equivalent). This space is linear and complete, hence Banach.

When considering only differentiable paths, e.g. the set $P(M, x_0) \cap C^1([0, 1], M)$, we will simply write $P(M, x_0) \cap C^1$.

Recall that the quotient of the path space $P(M, x_0)$ by the equivalence relation between paths homotopic with fixed ends, with the induced quotient topology, is exactly the universal covering space \tilde{M} .

By differentiating with respect to the time parameter t , we now transport the study of $P(G, e) \cap C^1$ to the study of the path space $P(\mathfrak{g})$ of the Lie algebra of G . The main question here is the following: given a path $\xi(t)$ in \mathfrak{g} , is it always possible to find a path $\gamma(t)$ in G starting at $e \in G$ whose velocity vector is $\xi(t)$, i.e. $(R_{\gamma(t)})_* \dot{\gamma}(t) = \xi(t)$. Of course, such a path is guaranteed to exist on some time interval $[0, \epsilon]$, but our goal is to show that in fact it exists on the entire interval $[0, 1]$. We formulate this as a homeomorphism because we also want to know that these solutions depend on ξ continuously and smoothly.

Lemma 11.9.3 ([33, Prop. 1.13.4]). *Let G be a connected Lie group with Lie algebra \mathfrak{g} . Then the mapping*

$$D = D^R : \quad \gamma \mapsto \left(t \mapsto \left(R_{\gamma(t)}^{-1} \right)_* \dot{\gamma}(t) \right) \quad (11.9.9)$$

is a homeomorphism from $P(G, e) \cap C^1$ onto $P(\mathfrak{g})$. For $\delta \in P(\mathfrak{g})$, let $A_\delta \in C^1([0, 1], L(\mathfrak{g}))$ be the solution A of the initial value problem

$$\dot{A}(t) = \text{ad}_{\delta(t)} \circ A(t), \quad A(0) = I = \text{id}_{\mathfrak{g}}. \quad (11.9.10)$$

Then, for all $\gamma, \gamma' \in P(G, e) \cap C^1$ and $t \in [0, 1]$:

$$D(\gamma \cdot \gamma')(t) = D\gamma(t) + \text{Ad}_{\gamma(t)} D\gamma'(t), \quad (11.9.11)$$

and

$$\text{Ad}_{\gamma(t)} = A_{D\gamma}(t). \quad (11.9.12)$$

Finally, denoting by $\Lambda(G)^\circ$ the space of all nullhomotopic based loops in G (with e being the basepoint)³², we have $D(\Lambda(G)^\circ \cap C^1) = P(\mathfrak{g})_0$, where $P(\mathfrak{g})_0$ is the set of the $\delta \in P(\mathfrak{g})$ satisfying the following conditions:

$$\exists \text{ smooth mapping: } s \mapsto \delta_s : [0, 1] \rightarrow P(\mathfrak{g}) \text{ with: } \delta_0 = 0, \delta_1 = \delta, \quad (11.9.13)$$

$$\int_0^1 A_{\delta_s}(t)^{-1} \partial_s \delta_s(t) dt = 0 \quad (s \in [0, 1]). \quad (11.9.14)$$

Proof. The statement that D is bijective means that for every continuous curve $\delta : [0, 1] \rightarrow \mathfrak{g}$ there is a unique C^1 solution curve $\gamma : [0, 1] \rightarrow G$ of the equation

$$\dot{\gamma}(t) = (\xi(t))_R(\gamma(t)), \quad (11.9.15)$$

where $(\xi(t))_R$ is the right-invariant vector field corresponding to $\xi(t) \in \mathfrak{g}$. The mapping $t \mapsto (\xi(t))_R$ is continuous from $[0, 1]$ to the space of analytic vector fields on G . Thus for each $s \in [0, 1]$ and $g \in G$, there is a unique maximal solution $\gamma(t, s, g)$ defined on the flow

domain $\mathcal{D}_{(s,g)} \subset [0, 1]$ such that $\gamma(s, s, x) = x$ and $s \in \mathcal{D}_{(s,g)}$. Defining the flow

$$\Phi^{t,s} : x \mapsto \gamma(t, s, x) \text{ from time } s \text{ to time } t, \text{ on } G^{t,s} = \{x \in G \mid t \in \mathcal{D}_{(s,g)}\}, \quad (11.9.16)$$

we also have the group property $\Phi^{t,s} \circ \Phi^{s,u} \subset \Phi^{t,u}$. Now, if γ satisfies the ODE above, then for any $g \in G$:

$$\partial_t(\gamma(t)g) = (R_g)_{*\gamma(t)} (R_{\gamma(t)})_{*e} \xi(t) = (R_{\gamma(t)g})_* \xi(t). \quad (11.9.17)$$

Hence $t \mapsto \gamma(t)g$ satisfies the same ODE. This shows that $G^{t,s} = G$ and $\Phi^{t,s}(g) = \Phi^{t,s}(e)g$ for all $g \in G$. The group property implies that the flow $\Phi^{t,s}$, constructed first only for small $|t - s|$, is globally defined on G and all $t, s \in [0, 1]$; in particular, the solution γ exists for all $s, t \in [0, 1]$. The continuous dependence of the solution on the vector fields implies that the inverse D^{-1} is continuous, so the first statement of the Lemma is proved.

The product rule for $D(\gamma \cdot \gamma')$ is easy to verify:

$$\partial_t(\gamma(t)\gamma'(t)) = (R_{\gamma'})_{*\gamma} \dot{\gamma} + (L_{\gamma})_{*\gamma'} \dot{\gamma}', \quad (11.9.18)$$

which then recombines into the claimed equality (11.9.11).

In addition we check

$$\partial_t \text{Ad}_{\gamma(t)} = \frac{d}{ds} \Big|_{s=0} \text{Ad}_{\gamma(t+s)\gamma(t)^{-1}} \text{Ad}_{\gamma(t)} = \text{ad}_{D\gamma(t)} \text{Ad}_{\gamma(t)}, \quad \text{Ad}_{\gamma(0)} = \text{Ad}_e = I, \quad (11.9.19)$$

and this shows that $A : t \mapsto \text{Ad}_{\gamma(t)}$ satisfies (11.9.10) with $\delta = D\gamma$.

If $u \mapsto \gamma_u$ is smooth $[0, 1] \rightarrow P(G, e) \cap C^1$ then, writing $\Phi_u^{t,s}$ for the flow from times s to time t , of the time-dependent vector field $(D\gamma_u(t))_R$, we have the variational formula from Lemma 11.9.1:

$$\partial_u \gamma_u(e) = \int_0^1 \left(T_{\gamma_u(s)} \Phi_u^{1,s} \right) (\partial_u (D\gamma_u(s))_R) \gamma_u(s) ds. \quad (11.9.20)$$

Now

$$\begin{aligned} T_{\gamma_u(s)} \Phi_u^{1,s} &= T_{\gamma_u(s)} \left(\Phi_u^{1,0} \circ (\Phi_u^{s,0})^{-1} \right) = \\ &= T_{\gamma_u(s)} \Phi_u^{1,0} \circ \left(T_e \Phi_u^{s,0} \right)^{-1} = T_e L_{\gamma_u(1)} \circ T_e L_{\gamma_u(s)}^{-1}, \end{aligned} \quad (11.9.21)$$

using the fact that the flow of a right-invariant field is performed by left translations. On the other hand,

$$((D\gamma_u)(s))_R(g) = (R_g)_* D\gamma_u(s), \quad (11.9.22)$$

so the variational formula takes the form

$$\partial_u \gamma(1) = (L_{\gamma_u(e)})_* \circ \int_0^1 \text{Ad}_{\gamma_u(s)}^{-1} \partial_u(D\gamma_u)(s) ds. \quad (11.9.23)$$

Then $u \mapsto \gamma_u(1)$ is constant iff $\int_0^1 \text{Ad}_{\gamma_u(s)}^{-1} \partial_u D\gamma_u(s) ds = 0$. Again using $\text{Ad}_{\gamma_u(s)} = A_{D\gamma_u}(s)$, we now obtain the last statement of this Lemma. ■

12 Lie Theory III: Analytic Structure

This section is devoted to the Dynkin/BCH formulas and their deep implications for the analytic structure of Lie groups. We have previously defined Lie subgroups as immersed subgroups, and in Theorem 11.4.2 we proved that all Lie groups are in fact weakly embedded. The main unanswered question is what topological conditions are sufficient for an arbitrary subgroup to be an immersed Lie subgroup, and in particular an embedded Lie subgroup. The main results of this section are as follows. The Closed Subgroup Theorem 12.4.2 says that all closed subgroups are embedded Lie subgroups. It implies, in particular, that every Lie group has a unique structure of a real analytic Lie group, and that all continuous homomorphisms between Lie groups are real analytic. The Analytic Subgroup Theorem 12.6.1 states the every Lie subalgebra of the Lie algebra of a Lie group generates a unique immersed Lie subgroup. Yamabe's Theorem 12.7.1 states that every path-connected subgroup of a Lie group is a Lie subgroup. Finally, the Initial Subgroup Theorem 12.7.2 states that every subgroup of a Lie group is (or rather can be turned into) a Lie subgroup. Therefore one of the main upshots of this sequence of results is:

$$\text{All subgroups of a Lie group are weakly embedded Lie subgroups.} \quad (12.0.1)$$

$$\text{Closed subgroups are embedded Lie subgroups.} \quad (12.0.2)$$

Another fundamental result in this section is Lie's Third Fundamental Theorem 12.5.6, which states that every finite-dimensional Lie algebra corresponds to a unique simply-connected Lie group. We prove only a local version of this theorem.

12.1 Product in logarithmic coordinates

The exponential map introduces a canonical (logarithmic) coordinate system in a neighborhood of the identity of a Lie group. This implies that the group structure, namely the multiplication map (the inversion map is trivial since $(e^X)^{-1} = e^{-X}$), can be expressed in these local coordinates. Namely let us introduce the multiplication map in logarithmic coordinates:

$$\mu(X, Y) = \ln(e^X e^Y), \quad X, Y \in \mathfrak{g}. \quad (12.1.1)$$

As we will show, μ has the – perhaps unsurprising but definitely not obvious – property that it can be expressed only in terms of repeated Lie brackets of X and Y , or, in other words, only in terms of the operators ad_X and ad_Y . Explicit expressions for $\mu(X, Y)$ in terms of these operators are collectively called the *Baker-Campbell-Hausdorff formula*, although the first complete and explicit version was obtained by Dynkin.

We will later show that every finite-dimensional real Lie algebra is the Lie algebra of some Lie group. Therefore it will be worthwhile to give an independent definition of $\mu(X, Y)$ on

an abstract Lie algebra. This is a bit tricky because in the absence of a Lie group there is no exponential map. The only object we have access to in the Lie algebra is a Lie bracket, or the adjoint operator ad . Therefore let us proceed in the following roundabout way.

For any finite-dimensional Lie algebra \mathfrak{g} , let \mathfrak{g}_e be the set of $X \in \mathfrak{g}$ such that the operator $f(\text{ad}_X) = (\text{e}^{\text{ad}_X} - I)(\text{ad}_X)^{-1}$ is invertible. The set \mathfrak{g}_e is an open neighborhood of 0 in \mathfrak{g} and, in view of the chain rule,

$$X \rightarrow \phi(\text{ad}_X) = (f(\text{ad}_X))^{-1} = \frac{\text{ad}_X}{\text{e}^{\text{ad}_X} - I} \quad (12.1.2)$$

is an analytic mapping $\mathfrak{g}_e \rightarrow L(\mathfrak{g})$ (the coefficients of its Taylor series are the Bernoulli numbers B_k). Let \mathfrak{g}_e^2 be the set of $(X, Y) \in \mathfrak{g} \times \mathfrak{g}_e$ such that the equation

$$\dot{Z}(t) = \phi(\text{ad}_{Z(t)}), \quad Z(0) = Y, \quad (12.1.3)$$

has a solution $Z(t)$ defined for all $t \in [0, 1]$. We now define

$$\mu(X, Y) = Z(1) \quad ((X, Y) \in \mathfrak{g}_e^2) \quad (12.1.4)$$

and claim that this formula defines an analytic mapping which matches the one obtained from the Lie group.

Theorem 12.1.1 ([33, Thm. 1.6.1]). *The set \mathfrak{g}_e^2 is an open neighborhood of $(0, 0)$ in \mathfrak{g}^2 and the map $\mu : \mathfrak{g}_e^2 \rightarrow \mathfrak{g}$ defined above is real-analytic. If \mathfrak{g} is the Lie algebra of a Lie group G , with exponential mapping $e^\cdot = \exp : \mathfrak{g} \rightarrow G$, then*

$$\text{e}^X \text{e}^Y = \text{e}^{\mu(X, Y)} \quad ((X, Y) \in \mathfrak{g}_e^2). \quad (12.1.5)$$

If \mathfrak{g} is a complex vector space and $\text{ad} : \mathfrak{g}^2 \rightarrow \mathfrak{g}$ is complex bilinear, then μ is complex-analytic.

Proof. The fact that \mathfrak{g}_e^2 is open and μ is real- (or complex-)analytic on \mathfrak{g}_e^2 follows from the fact that the map $(X, Z) \mapsto (f(\text{ad}_Z))^{-1}(X)$ is real-, or complex-, analytic: $\mathfrak{g} \times \mathfrak{g}_e \rightarrow \mathfrak{g}$, respectively, combined with the analytic dependence on initial values and parameters of solutions of analytic vector fields depending analytically on parameters. Now the formula for the differential of \exp combined with the equation for $Z(t)$ gives

$$\partial_t \text{e}^{Z(t)} = (T_{Z(t)} \exp) \dot{Z}(t) = (T_e R_{\text{e}^{Z(t)}})(X) = X_R \left(\text{e}^{Z(t)} \right). \quad (12.1.6)$$

In view of our previous result showing that the flow of a right-invariant vector field is performed by left translations, we have

$$\text{e}^{Z(t)} = \text{e}^{tX} \text{e}^{Z(0)} = \text{e}^{tX} \text{e}^Y, \quad (12.1.7)$$

for all t in the interval in the definition of $Z(t)$. Taking $t = 1$ gives the result. ■

Remark 12.1.2. This result can be used to show that the differentiable structure of every C^2 Lie group contains a real- or complex-analytic Lie group structure. Specifically such an atlas consists of the left translations $L_g U_0$ of a neighborhood U_0 on which \exp is invertible and μ is defined, and the chart maps $h \mapsto \ln(g^{-1}h)$ (that is, $\ln \circ L_g^{-1}$). Moreover, this structure is unique in the sense that any other analytic structure that is C^1 -diffeomorphic to the given one is actually analytically diffeomorphic to it. For a full proof see [33, Thm. 1.6.3, Prop. 1.6.4].

Remark 12.1.3. The fundamental meaning of this result is that the entire structure of the Lie group, only up to some elements of its global topology (namely the fundamental group), is fully preserved in its Lie algebra. This is one of the milestones on the way to the profound result that every abstract finite-dimensional Lie algebra is actually the Lie algebra of some Lie group (Lie's Third Fundamental Theorem).

12.2 Dynkin's formula

From the definition of $Z(t)$ in the last section combined with $\mathbf{Ad} \circ \exp = \exp \circ \text{ad}$, we get

$$e^{\text{ad}Z(t)} = e^{t\text{ad}_X} \circ e^{\text{ad}_Y}. \quad (12.2.1)$$

This can also be verified without reference to the Lie group by applying ad to the differential equation (12.1.3) defining $Z(t)$ and then reading (12.1.7) with X, Y , and Z replaced by ad_X, ad_Y , and ad_Z , respectively (which can always be done by sticking both sides into \mathbf{Ad} and using $\mathbf{Ad} \circ \exp = \exp \circ \text{ad}$). Now:

$$\begin{aligned} \dot{Z}(t) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^{\text{ad}Z(t)} - I)^k (X) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^{t\text{ad}_X} \circ e^{\text{ad}_Y} - I)^k (X) = \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \left(\sum_{l,m \geq 0, l+m>0} t^l \frac{(\text{ad}_X)^l}{l!} \circ \frac{(\text{ad}_Y)^m}{m!} \right)^k (X) = \\ &= \sum_{k=0}^{\infty} \sum_{l_1, \dots, l_k} t^{l_1+\dots+l_k} \left(\frac{(\text{ad}_X)^{l_1}}{l_1!} \circ \frac{(\text{ad}_Y)^{m_1}}{m_1!} \circ \dots \circ \frac{(\text{ad}_X)^{l_k}}{l_k!} \circ \frac{(\text{ad}_Y)^{m_k}}{m_k!} \right) (X), \end{aligned} \quad (12.2.2)$$

where the sum is taken over all $l_1, \dots, l_k, m_1, \dots, m_k \geq 0$ such that $l_j + m_j > 0$, for all j . Integrating this over t from 0 to 1 and using $Z(0) = Y$, we find *Dynkin's formula*:

$$\begin{aligned} \mu(X, Y) &= X + Y + \\ &+ \sum_{k=1}^{\infty} \frac{(-1)^k}{k+1} \sum_{l_1, \dots, l_k} \frac{1}{l_1 + \dots + l_k + 1} \left(\frac{(\text{ad}_X)^{l_1}}{l_1!} \circ \frac{(\text{ad}_Y)^{m_1}}{m_1!} \circ \dots \circ \frac{(\text{ad}_X)^{l_k}}{l_k!} \circ \frac{(\text{ad}_Y)^{m_k}}{m_k!} \right) (X). \end{aligned} \quad (12.2.3)$$

Remark 12.2.1. Although we obtained this series as a formal one, it actually always converges within some nonzero radius of the origin (assuming the Lie algebra is finite-dimensional). This follows from the analyticity of the exponential map, but also can be seen directly from the factorial-speed decay of the coefficients of the series. In fact, the BCH series is often derived purely algebraically and then used to *define* the analytic structure on the Lie group.

Recall that an element a of a ring is called nilpotent if $a^n = 0$ for some n and unipotent if $(a - 1)$ is nilpotent.

Corollary 12.2.2 ([37, Cor. 5.16]). *Let N be a connected Lie subgroup of $\mathrm{GL}_n(\mathbb{C})$ whose Lie algebra \mathfrak{n} lies in the set of strictly upper triangular matrices. Then the map $\exp : \mathfrak{n} \rightarrow N$ is bijective.*

Proof. It is a simple exercise to see that $[X_n, \dots, X_3, X_2, X_1] = [X_n, [\dots, [X_3, [X_2, X_1]] \dots]] = 0$ for any strictly upper triangular matrices $X, X_i \in \mathfrak{gl}_n(\mathbb{C})$ and that e^X is polynomial in X (because all strictly triangular matrices are nilpotent). In particular, for $X, Y \in \mathfrak{n}$ near the origin, Dynkin's Formula gives a polynomial expression for $Z \in \mathfrak{n}$ solving $e^X e^Y = e^Z$. Since both sides of this expression are polynomials in X and Y that agree on a neighborhood, they agree everywhere. Since the formula for Z involves only the Lie algebra structure of \mathfrak{n} , Z remains in \mathfrak{n} for any $X, Y \in \mathfrak{n}$. In other words, $(\exp \mathfrak{n})^2 \subset \exp \mathfrak{n}$. Since $\exp \mathfrak{n}$ generates all of N by Proposition 6.3.6, this shows that $\exp \mathfrak{n} = N$. This proves surjectivity.

The image of \exp , i.e. N , clearly consists of unipotent matrices: $e^X - 1$ is a polynomial in X with no constant term, therefore it is nilpotent just as X is. Bijectivity of \exp then follows from the fact that the inverse is given by $a \mapsto \ln(I + (a - I)) \in \mathfrak{n}$, which is also a terminating series due to the fact that $(a - I)$ is nilpotent. ■

Remark 12.2.3. A Lie algebra is called nilpotent if there exists an $n \in \mathbb{N}$ such that $\mathrm{ad}_{X_1} \mathrm{ad}_{X_2} \cdots \mathrm{ad}_{X_n} = 0$ for any sequence $X_1, \dots, X_n \in \mathfrak{g}$. A fundamental theorem of Engel states that a finite-dimensional Lie algebra is nilpotent iff each adjoint operator is nilpotent, i.e. $\mathrm{ad}_X^n = 0$ for all X . A Lie group G is called nilpotent if $\mathrm{Ad}_{g_1} \cdots \mathrm{Ad}_{g_n} g = e$ for all $g_1, \dots, g_n, g \in G$. Assuming G is connected, this is equivalent to its Lie algebra being nilpotent. The above Corollary is then a special case of a general statement that the exponential map is bijective for any *nilpotent* Lie algebra, and thus any nilpotent Lie group is diffeomorphic to a vector space.

Nilpotent Lie algebras and groups are a very important class because all of the group structure is polynomial, making it “almost abelian”.

12.3 Geometric BCH formula

In this section, let $\psi(z)$ be the function defined in the complex plane by

$$\psi(z) = \frac{z \ln z}{z - 1}, \quad (12.3.1)$$

or by its Taylor series at $z = 1$ (which converges in a disk of radius 1)

$$\psi(1 + u) = (1 + u) \frac{\ln(1 + u)}{u} = 1 + \frac{u}{2} - \frac{u^2}{6} + \dots . \quad (12.3.2)$$

Theorem 12.3.1. *The geometric Baker-Campbell-Hausdorff formula states that, with $\psi(z)$ defined by the above formal series, the group product in logarithmic coordinates can be expressed as*

$$\begin{aligned} \mu(X, Y) = \ln(e^X e^Y) &= Y + \int_0^1 \psi\left(e^{-\text{ad}_Y} e^{-t\text{ad}_X}\right) X dt = \\ &= X + \int_0^1 \psi\left(e^{\text{ad}_X} e^{t\text{ad}_Y}\right) Y dt. \end{aligned} \quad (12.3.3)$$

This means that we are meant to substitute

$$u = e^{\text{ad}_X} e^{t\text{ad}_Y} - I \quad (12.3.4)$$

into the formal series for $\psi(1 + u)$, apply each term to Y and then integrate. In doing this, we can ignore all terms where a factor of ad_Y shows up on the right since $\text{ad}_Y Y = 0$. To obtain the expansion through terms of degree three, we need only retain quadratic and lower terms in u :

$$u = \text{ad}_X + \frac{1}{2}(\text{ad}_X)^2 + t\text{ad}_Y + \frac{t^2}{2}(\text{ad}_Y)^2 + \dots , \quad (12.3.5)$$

$$u^2 = (\text{ad}_X)^2 + t\text{ad}_Y \text{ad}_X + \dots , \quad (12.3.6)$$

$$\int_0^1 \left(1 + \frac{u}{2} - \frac{u^2}{6}\right) dt = 1 + \frac{1}{2}\text{ad}_X + \frac{1}{12}(\text{ad}_X)^2 - \frac{1}{12}\text{ad}_Y \text{ad}_X + \dots , \quad (12.3.7)$$

where the dots indicate either higher order terms or terms with ad_Y occurring on the right. So the cubic Taylor expansion of $\mu(X, Y)$ reads

$$\mu(X, Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \dots , \quad (12.3.8)$$

agreeing with our previous results, including Dynkin's formula.

Proof. Our direct aim will be to prove the differential version of the formula:

$$\frac{d}{dt} \ln \left(e^{tX} e^Y \right) = \psi \left(e^{-\text{ad}Y} e^{-t\text{ad}X} \right) X. \quad (12.3.9)$$

Since $\ln e^{tX} e^Y = Y$ when $t = 0$, integrating the above formula from 0 to 1 will prove the geometric BCH formula. Let us define the curve γ by

$$\gamma(t) = \ln e^{tX} e^Y. \quad (12.3.10)$$

Then

$$\frac{d}{dt} e^{\gamma(t)} = (R_{e^{\gamma(t)}})_* X, \quad (12.3.11)$$

so

$$(R_{e^{-\gamma(t)}})_* \frac{d}{dt} e^{\gamma(t)} = X. \quad (12.3.12)$$

By applying **Ad** we also have

$$\text{ad}_\gamma(t) = \ln \left(e^{t\text{ad}X} e^{\text{ad}Y} \right). \quad (12.3.13)$$

Let us apply the Maurer-Cartan equation (11.1.12) to

$$g(t, s) = e^{s\gamma(t)}, \quad (12.3.14)$$

and get

$$\partial_s \xi - \gamma'(t) - [\gamma(t), \xi] = 0, \quad (12.3.15)$$

where

$$\xi(t, s) = (R_{e^{-s\gamma(t)}})_* \partial_t e^{s\gamma(t)}. \quad (12.3.16)$$

For fixed t this is a differential equation in s :

$$\partial_s \xi = \text{ad}_\gamma \xi + \gamma', \quad \xi(0) = 0, \quad (12.3.17)$$

where $\gamma = \gamma(t)$ and $\gamma' = \gamma'(t)$. If we expand $\xi(t, s)$ as a formal power series in s ,

$$\xi(t, s) = a_1 s + a_2 s^2 + a_3 s^3 + \dots, \quad (12.3.18)$$

and compare coefficients in the differential equation, we find $a_1 = \gamma'$ and

$$n a_n = \text{ad}_\gamma a_{n-1}, \quad (12.3.19)$$

or

$$\begin{aligned}\xi(t, s) &= s\gamma'(t) + \frac{s}{2}\text{ad}_{\gamma(t)}\gamma'(t) + \cdots + \frac{s^n}{n!}\text{ad}_{\gamma(t)}^{n-1}\gamma'(t) + \cdots = \\ &= \sum_{n=1}^{\infty} \frac{s^n}{n!}(\text{ad}_{\gamma})^{n-1}\gamma'(t).\end{aligned}\quad (12.3.20)$$

With the earlier notation $\phi(z) = (\mathrm{e}^z - 1)/z$, we have then at $s = 1$

$$(R_{\mathrm{e}^{-\gamma(t)}})_* \partial_t \mathrm{e}^{\gamma(t)} = \xi(t, 1) = \phi(\text{ad}_{\gamma(t)})\gamma'(t). \quad (12.3.21)$$

Now to the proof of the BCH formula. Suppose that X, Y are chosen small enough to be within the domain of $\mu(X, Y)$, so that $\gamma(t) = \mu(tX, Y)$ is defined for all $|t| \leq 1$. Then, by (12.3.13) and (12.3.12), (12.3.21) becomes

$$\phi\left(\ln \mathrm{e}^{t\text{ad}X} \mathrm{e}^{\text{ad}Y}\right)\gamma'(t) = X. \quad (12.3.22)$$

Now for $|z - 1| < 1$,

$$\phi(-\ln z) = \frac{z - 1}{z \ln z} = \psi(z)^{-1}, \quad (12.3.23)$$

thus

$$\gamma'(t) = \psi\left(\mathrm{e}^{-\text{ad}Y} \mathrm{e}^{-t\text{ad}X}\right)X. \quad (12.3.24)$$

This finishes the proof. ■

12.4 Closed subgroup theorem

Recall that in Theorem 6.3.12 we stated without proof that a Lie subgroup (immersed by definition) is embedded iff it is closed. Here we will use the exponential map to prove a much stronger statement that any closed subgroup of a Lie group is actually embedded.

Lemma 12.4.1 ([16, Lem. 5.6.7]). *If $\mathfrak{g} = \mathfrak{m}_1 \oplus \mathfrak{m}_2$, then there exist open neighborhoods U_i of the origin in \mathfrak{m}_i , $i = 1, 2$, and V of $e \in G$ such that the mapping*

$$\varphi : \mathfrak{m}_1 \times \mathfrak{m}_2 \rightarrow G, \quad (X_1, X_2) \mapsto \mathrm{e}^{X_1} \mathrm{e}^{X_2}, \quad (12.4.1)$$

restricts to a diffeomorphism $U_1 \times U_2 \rightarrow V$.

Proof. The differential $\varphi_{*0} : \mathfrak{m}_1 \times \mathfrak{m}_2 \rightarrow \mathfrak{g}$ is given by $(X_1, X_2) \mapsto X_1 + X_2$. Since \mathfrak{m}_1 and \mathfrak{m}_2 are complementary, it is bijective. Hence, the assertion follows from the InMT. ■

Theorem 12.4.2 (Closed Subgroup Theorem [16, Thm. 5.6.8]). *If G is a Lie group and $H \subset G$ is a subgroup that is also a closed subset of G , then H is an embedded Lie subgroup. The corresponding Lie subalgebra of $\mathfrak{g} = \text{Lie } G$ is*

$$\mathfrak{h} = \{X \in \mathfrak{g} : e^{tX} \in H \text{ for all } t \in \mathbb{R}\}. \quad (12.4.2)$$

Proof. By Proposition 6.3.2 it suffices to show that H is an embedded submanifold. We begin by identifying a subspace of $\text{Lie } G$ with what will turn out to be the Lie algebra of H .

Let $\mathfrak{g} = \text{Lie } G$ and define \mathfrak{h} as in the statement of the Theorem. We need to show that \mathfrak{h} is a linear subspace of \mathfrak{g} . It is obvious from the definition that \mathfrak{h} is closed under scalar multiplication: if $X \in \mathfrak{h}$, then $tX \in \mathfrak{h}$ for all $t \in \mathbb{R}$. Suppose $X, Y \in \mathfrak{h}$, and let $t \in \mathbb{R}$ be arbitrary. Then $\exp(tX/n)$ and $\exp(tY/n)$ are in H for any $n \in \mathbb{N}$, and since H is a closed subgroup, Corollary 11.5.11 implies that $\exp t(X + Y) \in H$. Thus $X + Y \in \mathfrak{h}$ and \mathfrak{h} is a subspace.

Next equip H with the subspace topology from G . Choose a neighborhood V of $e \in G$ such that $\ln = \exp^{-1} : V \rightarrow \mathfrak{g}$ is defined and smooth and hence a local chart. We claim that V can be adjusted so that

$$\ln(V \cap H) = \ln(V) \cap \mathfrak{h}. \quad (12.4.3)$$

If so, then the family consisting of local slice charts $\{(L_h(V), \ln \circ L_h^{-1})\}_{h \in H}$ satisfies the definition of a smooth structure on H and turns it into an embedded submanifold of G . Then H is also a Lie group because restricting the multiplication and inversion maps in range to an embedded submanifold keeps them smooth. The corresponding Lie subalgebra of \mathfrak{g} contains \mathfrak{h} and by (12.4.3) has the same dimension as \mathfrak{h} , hence coincides with \mathfrak{h} .

To show (12.4.3) assume, on the contrary, that V cannot be chosen so that (12.4.3) holds. Since $\exp(\mathfrak{h}) \subset H$, this means that $V \cap \exp(\mathfrak{h})$ is properly contained in $V \cap H$, for arbitrarily small V . Thus there is a sequence $(h_n) \in H$ converging to e such that $h_n \notin \exp(\mathfrak{h})$ for all n . Choose a subspace $\mathfrak{m} \subset \mathfrak{g}$ complementary to \mathfrak{h} . By Lemma 12.4.1, for large enough n , h_n defines unique sequences $(X_n) \in \mathfrak{h}$ and $(Y_n) \in \mathfrak{m}$ by $h_n = e^{X_n}e^{Y_n}$ and both of these sequences converge to the origin. Since H is a subgroup, we obtain $e^{Y_n} \in H$. Since $h_n \notin \exp(\mathfrak{h})$, we conclude $Y_n \neq 0$. Now, choose a norm $\|\cdot\|$ on \mathfrak{m} . Since the corresponding unit sphere in \mathfrak{m} is compact, by possibly removing some members we may assume that the sequence $(Y_n/\|Y_n\|)$ converges to some $Y \in \mathfrak{m}$ with $\|Y\| = 1$.

Now pick any $t \in \mathbb{R}$ and define the sequence of integer parts $k_n = \lfloor \frac{t}{\|Y_n\|} \rfloor$. Then $k_n\|Y_n\| \rightarrow t$ as $n \rightarrow \infty$ and hence

$$\exp tY = \lim_{n \rightarrow \infty} \exp \left(k_n\|Y_n\| \cdot \frac{Y_n}{\|Y_n\|} \right) = \lim_{n \rightarrow \infty} (\exp Y_n)^{k_n} \in H, \quad (12.4.4)$$

as H is closed. This means that $\exp tY \in H$ for all $t \in \mathbb{R}$ and hence $Y \in \mathfrak{h}$, which is a contradiction. ■

We summarize all results about closed subgroups of Lie groups in the following Corollary.

Corollary 12.4.3. *If G is a Lie group and H is a subgroup of G , then TFAE:*

1. H is closed in G .
2. H is an embedded submanifold of G .
3. H is an embedded Lie subgroup of G .

Example 12.4.4. The kernel of any Lie group homomorphism $\Phi : G \rightarrow H$ is a subgroup that is closed (by continuity) and hence an embedded Lie subgroup of G . From naturality of the exponential map, we easily conclude that the Lie algebra of $\ker \Phi$ is given by $\ker \Phi_{*e}$.

Example 12.4.5. In view of the closed subgroup theorem, to show that classical groups defined as subsets of $\mathrm{GL}_n(\mathbb{R})$ (which we already know to be a Lie group) it suffices to show that they are closed. This smooth structure will coincide with the one provided by the Level Set Theorem 3.9.7.

Corollary 12.4.6 ([33, Prop. 1.10.8]). *Assume G, H are Lie groups and H has only countably many components. Then every homomorphism of groups $\Phi : G \rightarrow H$ with a closed graph $\subset G \times H$ is analytic, that is, Φ is a homomorphism of Lie groups.*

Further, if Φ is a continuous isomorphism of groups, then it is an isomorphism of Lie groups (i.e. both Φ and Φ^{-1} are analytic).

Proof. Write $F = \mathrm{graph}(\Phi)$. Since Φ is a homomorphism, F is a subgroup of $G \times H$, and because it is closed, F is a Lie subgroup in view of the Closed Subgroup Theorem. Let π_1, π_2 be the projections $F \rightarrow G$ and $F \rightarrow H$, respectively. Since the dimension of $T_f\pi_1$ is constant (as a function of $f \in F$), the mapping π_1 is, locally in F , a fibration with fiber dimension equal to the dimension of $\ker T_f\pi_1$. However, π_1 is invertible, so the fibers consist of single points and we conclude that π_1 is an immersion.

Because $\Phi \circ L_g = L_{\Phi(g)} \circ \Phi$ for any $g \in G$, it is sufficient to prove analyticity for Φ is a neighborhood of e in G , or in G_0 , so we may assume that G is connected. It follows that $G \times H$, and therefore F , is a countable union of compact subsets, because F is closed in $G \times H$.

We conclude that if $\dim F < \dim G$ then $\pi_1(F)$ has measure zero in G , in contradiction with $\pi_1(F) = G$. So $\dim F \geq \dim G$ and π_1 is actually a local diffeomorphism, and hence a diffeomorphism because π_1 is invertible. It finally follows that $\Phi = \pi_2 \circ \pi_1^{-1}$ is analytic $G \rightarrow H$.

The result about isomorphisms follows immediately because under this assumption G must also have countably many components and we can apply the above to Φ^{-1} . ■

Remark 12.4.7. Note that without any assumption on H this Corollary is false: take $\dim G > 0$ and $H = G$ with discrete topology, and $\Phi = \text{id}_G$. The graph of Φ is closed because $\Phi^{-1} : H \rightarrow G$ is continuous.

From here, the next Corollary is trivial.

Corollary 12.4.8. *Every continuous homomorphism between Lie groups is analytic.*

Proof. A continuous mapping maps the identity component into the identity component, so we may assume both groups to be connected. Also a continuous map has a closed graph. ■

Corollary 12.4.9. *A topological group has at most one Lie group structure.*

Proof. If there are two Lie group structures, then the identity map provides a continuous and hence analytic isomorphism between them. ■

Due to this, we may think of Lie groups as a special class of topological groups. We may therefore ask which subgroups of a Lie group G are Lie groups with respect to the subspace topology.

Proposition 12.4.10 ([40, Prop. 9.3.9]). *A subgroup H of a Lie group is a Lie group with respect to the induced topology and smooth structure iff it is closed.*

Proof. If H is closed, then the Closed Subgroup Theorem 12.4.2 implies that H is an embedded Lie subgroup and hence agrees with the induced topology.

Conversely, suppose that H is a Lie group with respect to the induced topology and smooth structure. Since the inclusion $i : H \hookrightarrow G$ is continuous, Corollary 12.4.8 implies that i is actually smooth, and since it's injective, so is $i_* : \mathfrak{h} \rightarrow \mathfrak{g}$.

Let $V \subset \mathfrak{g}$ be an open convex neighborhood of the origin for which $\exp_G|_V$ is a diffeomorphism onto an open subset of G . Since i_* is continuous, there exists a neighborhood of the origin $U \subset \mathfrak{h}$ such that $\exp_H|_U$ is a diffeomorphism onto an open subset of H and $i_*(U) \subset V$. Since H is a topological subgroup of G , there exists an open subset $U_{\mathfrak{g}} \subset V$ with

$$\exp_G(U_{\mathfrak{g}}) \cap H = i(\exp_H(U)) = \exp_G(i_*U). \quad (12.4.5)$$

Since $i_*(U)$ is locally closed in $i(G)$, it now follows that H is locally closed in G , hence closed (by translation). ■

Remark 12.4.11. In Theorem 11.4.2 we saw that each immersed Lie subgroup is in fact weakly embedded. By Theorem 3.9.15, each weakly embedded manifold has a unique topology and smooth structure under which it's still immersed. This implies that, assuming a subset of G admits any structure of an immersed Lie group (this includes a topology, a smooth structure, and a compatible Lie group structure), then that structure is unique (you can also find this stated as a separate elementary lemma e.g. in [40, Lem. 9.6.12]). Proposition 12.4.10 further shows that if this structure agrees with the one induced from G , then the subgroup must be embedded/closed.

We will finish this Part with another fundamental theorem (Initial Subgroup Theorem 12.7.2) stating that *any* subgroup of G can in fact be turned into an immersed, hence weakly embedded, Lie subgroup. This makes the concept of a Lie subgroup indistinguishable from a set-theoretic subgroup and is the reason why some authors only consider closed subgroups to be worthy of the name “Lie subgroup”.

Example 12.4.12. Let $G = (\mathbb{R}, +)$ and $H = \mathrm{GL}(V)$ with V a finite-dimensional vector space. Let $h : G \rightarrow H$ be such that:

- (i) $h(t+s) = h(t)h(s)$ for all $t, s \in \mathbb{R}$;
- (ii) if $t_n \rightarrow 0 \in \mathbb{R}$ and $h(t_n) \rightarrow A \in \mathrm{GL}(V)$, then $A = I$.

Then the conclusion is that there is some $B \in \mathrm{End}(V)$ such that $h(t) = e^{tB}$ for all $t \in \mathbb{R}$. The assumption (ii) can even be further weakened because one can show that discontinuous homomorphisms $(\mathbb{R}, +) \rightarrow \mathrm{GL}(V)$ must behave rather wildly. One of the strongest theorems of this type exists in the realm of unitary operators Hilbert spaces – **Stone’s theorem on strongly continuous one-parameter unitary groups**.

12.5 Lie’s fundamental theorems

Definition 12.5.1 (Local Lie group). A local Lie group is a triple (U, μ, ι) where U is an open neighborhood of the origin in a finite-dimensional vector space \mathfrak{g} , and two real-analytic maps $\mu : U \times U \rightarrow \mathfrak{g}$, $\iota : U \rightarrow \mathfrak{g}$ such that for $X, Y, Z \in \mathfrak{g}$ sufficiently close to 0:

$$\mu(X, 0) = X = \mu(0, X), \tag{12.5.1}$$

$$\mu(X, \iota(Y)) = \mu(\iota(X), Y) = 0, \quad \mu(\mu(X, Y), Z) = \mu(X, \mu(Y, Z)). \tag{12.5.2}$$

Two local Lie groups (U, μ, ι) and (U', μ', ι') are considered isomorphic if there are open neighborhoods $U_{00} \subset U_0 \subset U$ of 0 and a diffeomorphism $\phi : U_0 \rightarrow U'_0 \subset U'$ with $0 \in U'_0$ such that for any $X, Y \in U_{00}$, $\mu(X, Y) \in U_0$, $\phi(\mu(X, Y)) = \mu'(\phi(X), \phi(Y))$ and $\iota(X) \in U_0$, $\phi(\iota(X)) = \iota'(\phi(X))$ whenever $X, Y \in U_{00}$. An isomorphism class of triples (U, μ, ι) is called a *Lie group germ*.

It is almost obvious that the definition of the exponential map together with its properties

have local versions that are valid for any local Lie group. The first big result of this section will be the converse: every Lie algebra generates a local Lie group. Less obvious is the global version of this statement, known as Lie's Third Fundamental Theorem: every Lie algebra is the Lie algebra of a unique simply-connected Lie group. The Lie group will be reconstructed as the group of flows of certain vector fields on the Lie algebra. Let us now show, in several steps, that these flows form a local Lie group.

Lemma 12.5.2 ([33, Lem. 1.8.2]). *For each $X \in \mathfrak{g}$ define the vector field $X^R \in \mathfrak{X}(\mathfrak{g}_e)$, where \mathfrak{g}_e was defined in Section 12.1:*

$$X^R(Z) = \frac{\text{ad}_Z}{e^{\text{ad}_Z} - I} X \quad (Z \in \mathfrak{g}_e). \quad (12.5.3)$$

Then for each $X, Y \in \mathfrak{g}$, the map $X \mapsto X^R$ is an injective Lie algebra anti-homomorphism $\mathfrak{g} \rightarrow \mathfrak{X}(\mathfrak{g}_e)$, i.e. $[X^R, Y^R] = -([X, Y])^R$.

Proof. Recall that for vector fields $v, w \in \mathfrak{X}(U \subset \mathbb{R}^n)$, if we denote their differentials as a point by $Dv(x) \in \text{End}(\mathbb{R}^n)$, then

$$[v, w](x) = Dw(x) \cdot v(x) - Dv(x) \cdot w(x). \quad (12.5.4)$$

So in order to compute $[X^R, Y^R](Z)$, we start by differentiating 11.6.12:

$$\int_0^1 e^{t\text{ad}_Z} dt \left(X^R(Z) \right) = X \quad (Z \in \mathfrak{g}_e), \quad (12.5.5)$$

with respect to Z , in the direction of $Y^R(Z)$. We get, using the formula for the derivative of the exponential from Theorem 11.6.1,

$$\begin{aligned} \int_0^1 \left(\int_0^1 e^{s t \text{ad}_{\text{ad}_Z}} ds \circ t \text{ad}_{Y^R(Z)} \right) \circ e^{t \text{ad}_Z} dt \left(X^R(Z) \right) + \\ + \int_0^1 e^{t \text{ad}_Z} dt \circ D X^R(Z) \left(Y^R(Z) \right) = 0. \end{aligned} \quad (12.5.6)$$

Since ad is a Lie algebra homomorphism $\mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ (this is equivalent to the Jacobi identity), one has:

$$e^{\text{ad}_{\text{ad}_X}} \circ \text{ad}_Y = \text{ad}_{e^{\text{ad}_X} Y} \quad (X, Y \in \mathfrak{g}). \quad (12.5.7)$$

Reading this with X , and Y , replaced by $s t Z$, and $Y^R(Z)$, respectively, and substituting $s t = u$, we can recognize the first term in the left hand size of (12.5.6) as

$$\iint_{0 \leq u \leq t \leq 1} \left[e^{u \text{ad}_Z} \left(Y^R(Z) \right), e^{t \text{ad}_Z} \left(X^R(Z) \right) \right] du dt. \quad (12.5.8)$$

Antisymmetrizing this formula over X and Y we get, also using the antisymmetry of the Lie bracket, the quantity

$$\begin{aligned} - \int_0^1 \int_0^1 & \left[e^{u \text{ad}_Z} \left(X^R(Z) \right), e^{t \text{ad}_Z} \left(Y^R(Z) \right) \right] du dt = \\ & = - \left[\int_0^1 e^{u \text{ad}_Z} du \left(X^R(Z) \right), \int_0^1 e^{t \text{ad}_Z} dt \left(Y^R(Z) \right) \right] = -[X, Y]. \end{aligned} \quad (12.5.9)$$

So, if we antisymmetrize (12.5.6) over X and Y , we arrive, using (12.5.4) with $v = X^R$, $w = Y^R$ and $x = Z$, at:

$$-[X, Y] = \int_0^1 e^{t \text{ad}_Z} dt \left([X^R, Y^R](Z) \right), \quad (12.5.10)$$

which shows that $[X^R, Y^R] = -([X, Y])^R$ in view of the definition (12.5.3) applied to $([X, Y])^R$. The linearity of the map $X \mapsto X^R$ is obvious from the definition, and its injectivity follows from $X^R(0) = X$. ■

Remark 12.5.3. Note that if \mathfrak{g} were known to be the Lie algebra of a Lie group G , then we would have $X^R(Z) = (R_{e^{-Z}})_{*e^Z} \circ \exp_{*Z}(Z)$, which is nothing but the right-invariant vector field $X_R(g) = (R_g)_{*e^Z} X$ written in logarithmic coordinates $g = e^Z$.

If all vector fields X^R were complete on \mathfrak{g}_e , their flows would form a Lie group consisting of diffeomorphisms of \mathfrak{g}_e , and we would prove the global Lie's Third Fundamental Theorem, which states that every Lie algebra has a Lie group that it corresponds to. Unfortunately, these vector fields are not always complete, and the most we can claim at this point is that their flows generate a *local* Lie group.

Theorem 12.5.4 ([33, Thm. 1.8.3]). *Let M be a C^2 manifold and let $\mathfrak{g} \subset \mathfrak{X}^1(M)$ be a finite-dimensional Lie subalgebra of the algebra of C^1 vector fields on M . Then, for every relatively compact open subset V of M (i.e. such that the closure \bar{V} is compact), there is an open neighborhood U of $0 \in \mathfrak{g}$ such that:*

- (a) *For every $X \in U$, the set \mathcal{D}_1^X , the slice of the flow domain of X at time $t = 1$, contains V .*
- (b) *If $X, Y \in U$, then $F_1^X(V) \subset \mathcal{D}_1^X$ and $(F_1^X \circ F_1^Y)|_V = F_1^{\mu(X, Y)}|_V$.*
- (c) *$X \mapsto F_1^X|_V$ is injective on U .*

In particular, the collection of flows $\{F_1^X\}_{X \in \mathfrak{g}}$ defines a local Lie group with Lie algebra equal to \mathfrak{g} .

Proof. Since we are only interested in time $t = 1$, let us drop the subscript 1 and just write F^X, F^Y and \mathcal{D}^X . Assertion (a) follows from the local existence theorem for solutions of

ODE's, combined with the locally uniform estimates for the domains of existence with respect to the parameters on which the vector fields depend continuously. Similarly we obtain $F^Y(V) \subset \mathcal{D}^X$ for all $X, Y \in U$, by sufficiently shrinking U .

In order to prove (b), consider, for $X, Y \in \mathfrak{g}$, the vector field on M :

$$Y(t) = (F^{tX})_*^{-1} (\mathrm{e}^{t\mathrm{ad}X} Y); \quad (12.5.11)$$

here $\mathrm{e}^{t\mathrm{ad}X}(Y) \in \mathfrak{g}$ is computed in terms of the finite-dimensional Lie algebra structure of \mathfrak{g} and F^{tX} is the flow of the vector field X after time t . For any local diffeomorphism F and vector field v , the vector field F_*v is defined by:

$$(F_*v)(F(x)) = T_x F(v(x)) \text{ for } x \text{ in domain of } F. \quad (12.5.12)$$

F_*v can be interpreted as $\mathrm{Ad}_F v$ if the space of vector fields on M is viewed as the Lie algebra of the group of transformations, and then F^X can be interpreted as e^X . In order to avoid confusion we don't yet make these identifications here, but the following arguments will provide support for this point of view.

Differentiating (12.5.11) with respect to t , we get

$$\frac{dY}{dt}(t) = -[X, Y(t)] + (F^{tX})_*^{-1} [X, \mathrm{e}^{t\mathrm{ad}X} Y] = -[X, Y(t)] + [X, Y(t)] = 0, \quad (12.5.13)$$

where we used the naturality of Lie brackets and $(F^{tX})_*^{-1} X = X$. Because $Y(0) = Y$, the conclusion is that $Y(t) = Y$, or:

$$(F^{tX})_* Y = \mathrm{e}^{t\mathrm{ad}X} Y \text{ on } V_t X. \quad (12.5.14)$$

(This confirms the interpretation above, see Theorem 11.5.8(a).) Now consider $Z(t) = \mu(tX, Y)$. Then, using (11.9.2) for the computation of $\partial_t F^{Z(t)}(x)$, replacing ϵ and t by t and 1, respectively, we obtain:

$$\begin{aligned} \frac{d}{dt} (F^{-tX} \circ F^{Z(t)}(x)) &= -X (F^{-tX} \circ F^{Z(t)}(x)) + \\ &\quad + T_{F^{Z(t)}(x)} F^{-tX} \circ T_x F^{Z(t)} \circ \int_0^1 (F^{sZ(t)})_*^{-1} \frac{dZ}{dt}(s) ds (x). \end{aligned} \quad (12.5.15)$$

Now, combining (12.5.14) with (11.6.12) and (12.1.3), we get:

$$\begin{aligned} \int_0^1 (F^{sZ(t)})_*^{-1} \frac{dZ}{dt}(s) ds &= \int_0^1 \mathrm{e}^{-s\mathrm{ad}Z(t)} ds \cdot \frac{dZ}{dt}(t) = \\ &= \mathrm{e}^{-\mathrm{ad}Z(t)} X = (F^{Z(t)})_*^{-1} X. \end{aligned} \quad (12.5.16)$$

If we apply (12.5.12) with F and v replaced by $F^{Z(t)}$ and X , respectively, we see that the right hand side in (12.5.15) is equal to 0. Hence $F^{-tX} \circ F^{Z(t)} = F^{Z(0)} = F^Y$, or:

$$F^{Z(t)} = F^{tX} \circ F^Y, \quad (12.5.17)$$

which for $t = 1$ becomes the required identity in (b).

The last statement follows by an argument similar to that in the Closed Subgroup Theorem 12.4.2, because otherwise we would get sequences $X_j, Y_j \in \mathfrak{g}$ converging to 0 and sets $V_j \subset M$ such that $X_j \neq Y_j$, for all j , but $F^{X_j}|_{V_j} = F^{Y_j}|_{V_j}$, and such that, for each relatively compact subset $V \subset M$, there is a j_0 with $V \subset V_j, j \geq j_0$. Passing to $Z_j = \mu(X_j, -Y_j)$, we get a sequence in \mathfrak{g} , converging to 0, such that $Z_j \neq 0$ and $F^{Z_j}|_{V_j}$ is the identity on V_j , for a sequence V_j of subsets of M as above. Passing to a suitable subsequence, we may assume that $Z_j/|Z_j|$ converges in \mathfrak{g} to, say Z . Here $X \mapsto |X|$ denotes a norm in \mathfrak{g} ; clearly $|Z| = 1$, so in particular $Z \neq 0$. Let $t \geq 0$ and $m_j = \lfloor \frac{t}{|Z_j|} \rfloor$, the integer part of $t/|M_j|$. Because $|Z_j| \rightarrow 0$ we have that $m_j|Z_j| \rightarrow t$ as $j \rightarrow \infty$, so $m_j Z_j = (m_j|Z_j|)(Z_j/|Z_j|) \rightarrow tZ$.

On the other hand,

$$F^{m_j Z_j} = \left(F^{Z_j} \right)^{m_j} = \text{id}_{V_j}, \quad (12.5.18)$$

and therefore we conclude that $F^{tZ} = \text{id}_M$ for all $t \geq 0$. Differentiation with respect to t at $t = 0$ no yields $Z = 0$, a contradicton. ■

Theorem 12.5.5 (Lie's Third Fundamental Theorem [33, Thm. 1.8.1]). *For every finite-dimensional Lie algebra \mathfrak{g} there exists a local Lie group whose Lie algebra is \mathfrak{g} . Namely, an open neighborhood of 0 in \mathfrak{g} becomes a local Lie group, provided with μ defined in (12.1.4) as the local multiplication and $X \mapsto -X$ as the local inversion.*

Proof. The properties of μ are easily verified except for the associativity. Associativity follows from the associativity of compositions of flows as well as (b) and (c) of Theorem 12.5.4. ■

In fact, a global version of this theorem holds. We only provide the statement without proof. The proof relies on the theory of infinite-dimensional Lie groups, and the basic idea is to show that equation (11.9.11) defines a group structure on the path space $P(\mathfrak{g})$ which makes it a Banach Lie group, and the identity component $P(\mathfrak{g})_0$ into a closed normal subgroup. Then the quotient $P(\mathfrak{g})/P(\mathfrak{g})_0$ turns out to be a simply connected Lie group with Lie algebra isomorphic to \mathfrak{g} . Showing that $P(\mathfrak{g})_0$ is closed requires either a deep exploration of the topology of Lie groups (one needs to show that $\pi_2(G) = 0$ for any Lie group, or at least that the second cohomology group is trivial, as in [33]), or structure theory of Lie algebras (via Ado's theorem).

Theorem 12.5.6 (Global Lie's Third Fundamental Theorem [33, Thm. 1.14.3]). *For every finite-dimensional Lie algebra \mathfrak{g} there exists a simply connected Lie group \tilde{G} whose Lie algebra is equal to \mathfrak{g} . The restriction of the exponential mapping $\exp : \mathfrak{g} \rightarrow \tilde{G}$ to the center $\mathfrak{z} = Z(\mathfrak{g})$ induces an isomorphism from the vector group $(\mathfrak{z}, +)$ onto the identity component of the center $Z(\tilde{G})$.*

Remark 12.5.7. Lie's Third Fundamental Theorem, combined with Theorem 11.3.10, states that the Lie functor provides an equivalence of categories

$$\text{Lie} : \text{SLieGr} \rightarrow \text{LieAlg}_{\mathbb{R}}, \quad (12.5.19)$$

where SLieGr is the category of simply connected Lie groups.

One of the other two fundamental theorems have already been covered by us. Their original statements were about local groups, but since we won't use them, we provide the following alternative versions.

Lie's First Fundamental Theorem Lie groups G and H are locally isomorphic, iff their Lie algebras $\text{Lie } G$ and $\mathcal{L}H$ are isomorphic. This was shown in Proposition 6.10.13 and Theorem 11.3.10.

Lie's Second Fundamental Theorem There is a bijective correspondence between Lie subalgebras of $\text{Lie } G$ and connected Lie subgroups of G . This will be shown in the Analytic Subgroup Theorem 12.6.1.

Remark 12.5.8. Lie's Third Fundamental Theorem gives a one-to-one correspondence between local data (Lie algebra) and global nonlinear data (group multiplication). There are several explicit ways of constructing the global Lie group:

1. As a quotient of something simpler – e.g. the path space considered above. More concretely, given a simply connected lie group G , one can consider the space of all smooth paths from e to $g \in G$ and declare two paths to be equivalent iff there is a pair of functions $\xi(t, s), \eta(t, s) \in \mathfrak{g}$ such that the Maurer-Cartan equation (11.1.12) holds for it, $\partial_t \eta - \partial_s \xi = [\xi, \eta]$ with proper boundary conditions. This equation describes the vanishing of curvature along the surface swept out by the homotopy. Therefore G can be reconstructed as the space of paths in \mathfrak{g} starting at the origin modulo solutions of this nonlinear equation. Fortunately, the analysis reduces to first-order deformations.
2. As a “submanifold” of something simpler. Each Lie algebra \mathfrak{g} can be embedded into a matrix Lie algebra $\text{gl}_n(\mathbb{R})$. Then we have a map $G \rightarrow \text{GL}_n(\mathbb{R})$ with some kernel Γ and G is the universal covering group of $G/\Gamma \subset \text{GL}_n(\mathbb{R})$. Thus, at least locally, every element of the Lie group is a matrix. The problem with this approach is that G/Γ need

not be a manifold. Generally, for an immersed subgroup $H \subset G$, we have a foliation of G by the cosets of H . In particular, we have the distribution consisting of translated copies of \mathfrak{h} , i.e. $(R_g)_*\mathfrak{h}$. Thus the cosets can be reconstructed as integral manifolds of this integrable distribution.

Thus we can define the distribution $D_g = \mathfrak{g} \cdot g$ on $\mathrm{GL}_n(\mathbb{R})$, verify that it is involutive, and consider the integral manifold through $I \in \mathrm{GL}_n(\mathbb{R})$. Since the distribution is right-invariant, the integral manifold, which we can call G , is too. If $h \in G$, then $G \cdot h^{-1}$ is also integral and contains I , thus $G \cdot h^{-1} = G$, hence G is a connected group. If G is not simply connected, we can take its universal covering group instead. This is the essence of Ado's Theorem, which says that every Lie group can be realized as an immersed Lie subgroup of a matrix group GL_n . This is another strategy for proving the Third Fundamental Theorem.

Exercise 12.1. Verify that $U = (-\frac{1}{2}, \frac{1}{2})$ is a local Lie group with (abelian) multiplication and inversion given by, respectively:

$$\mu(x, y) = \frac{2xy - x - y}{xy - 1}, \quad \iota(x) = \frac{x}{2x - 1}. \quad (12.5.20)$$

Prove $\kappa : (-\infty, 1) \rightarrow \mathbb{R}$ with $\kappa(x) = x/(x - 1)$ satisfies $\kappa(x + y) = \mu(\kappa(x), \kappa(y))$ and $\kappa(-x) = \iota(\kappa(x))$ where defined. Deduce that μ and ι are the usual addition and inversion on \mathbb{R} in a local chart for \mathbb{R} .

Now use Theorem 12.5.6 to show the following general result. Let $U \subset \mathfrak{g}$ be a local Lie group with multiplication μ and inversion ι . Then there exists a simply connected Lie group G and a coordinate chart $\kappa : V \rightarrow U$ where V is a neighborhood of $e \in G$, such that:

$$\kappa(e) = 0, \quad \kappa(xy) = \mu(\kappa(x), \kappa(y)), \quad \kappa(x^{-1}) = \iota(\kappa(x)), \quad (12.5.21)$$

whenever $x, y \in V$.

12.6 Analytic subgroup theorem

In the Closed Subgroup Theorem 12.4.2 we characterized embedded Lie subgroups and described their Lie algebras. We now study the connection between general Lie subgroups and Lie subalgebras. Already from the simple example of one-dimensional subgroups of the torus \mathbb{T}^2 it is clear that the answer will be subtle: some Lie subalgebras will generate embedded subgroups, while generically they're only guaranteed to generate weakly embedded subgroups, that is, subgroups that become Lie groups only after adjusting their topology and smooth structure away from the ones induced from G . The connected subgroup generated by a subalgebra $\mathfrak{h} < \mathrm{Lie} G$ is often called an *analytic subgroup*, and the following important

theorem shows that it can always be made into an immersed Lie subgroup.

Theorem 12.6.1 (Analytic Subgroup Theorem [33, Thm. 1.10.3], [16, Prop. 5.6.5], [40, Thm. 9.4.8]). *Given a Lie group G , for each Lie subalgebra $\mathfrak{h} \subset \mathfrak{g} = \text{Lie } G$ there is a unique connected immersed Lie subgroup H of G whose Lie algebra is \mathfrak{h} . Moreover, the Lie algebra \mathfrak{h} satisfies*

$$\mathfrak{h} = \{X \in \mathfrak{g} : e^{tX} \in H \text{ for all } t \in \mathbb{R}\}. \quad (12.6.1)$$

Proof. One proof of the first statement, taken from [16, Prop. 5.6.5], is based on the Global Frobenius Theorem 9.5.13. The Lie subalgebra of $\mathfrak{X}_L(G)$ that consists of left-invariant vector fields corresponding to elements of \mathfrak{h} spans an involutive distribution on G . Then H can be defined as the maximal integral submanifold through $e \in G$. Now since the distribution is left-invariant, L_h^{-1} for $h \in H$ takes the integral manifold through $h \in H$ to the one through e , which is H . Therefore $L_h(H) = H$, and thus H is a subgroup. The multiplication on H is smooth by virtue of being the restriction of the smooth multiplication map $m : H \times H \rightarrow G$ in range to an weakly embedded submanifold of G (which an integral manifold is by Proposition 9.5.7).

Now we present the proof from [33, Thm. 1.10.3], which avoids the Frobenius Theorem and is instead based on the exponential map. For each $X \in \mathfrak{h}$, the vector field X^R on \mathfrak{g}_e defined in (12.5.3) is tangent to \mathfrak{h} in the sense that $X^R(Z) \in \mathfrak{h}$ for all $Z \in \mathfrak{g}_e \cap \mathfrak{h}$. This follows because $\text{ad}_Z \in L(\mathfrak{h})$ if $Z \in \mathfrak{h}$. Moreover \mathfrak{h} is a closed subspace of \mathfrak{g} as a finite-dimensional subspace, hence the power series in the right hand side of (12.5.3) converges to an element of \mathfrak{h} . (In the case of Banach Lie groups we have to make the additional assumption that \mathfrak{h} is a closed linear subspace.) If we solve (12.1.3) with $Z(0) = Y \in \mathfrak{h}$, regarded as an ODE in $\mathfrak{g}_e \cap \mathfrak{h}$, the solution will also be in $\mathfrak{g}_e \cap \mathfrak{h}$. However, it is also equal to the solution of the same ODE viewed in \mathfrak{g}_e , which is equal to $\ln(e^{tX}e^Y)$. So the conclusion is that $\ln(e^{tX}e^Y) \in \mathfrak{h}$ if $X, Y \in \mathfrak{g}_e \cap \mathfrak{h}$. Denote by H the subgroup generated by $\exp(\mathfrak{h})$, then the chart maps $\kappa^h|_H$, for $h \in H$, with $\kappa^g = \ln \circ L_g^{-1}$, form an atlas for H , mapping it into \mathfrak{h} . This can be checked to be an analytic atlas for H (see again [33, Thm. 1.6.3]). The inclusion $H \hookrightarrow G$ is an immersion because, in the charts κ^h , it is equal to the inclusion of open subsets of \mathfrak{h} into \mathfrak{g} . In view of Proposition 6.3.6, H is connected, and it is the only connected Lie group of G having \mathfrak{h} as its Lie algebra. The uniqueness of the structure of a real-analytic group on H with prescribed Lie algebra \mathfrak{h} is also very easy to show by a standard argument, see [33, Prop. 1.6.4].

Finally we need to show that if $e^{tX} \in H$ for all $t \in \mathbb{R}$, then $X \in \mathfrak{h}$. Let $V \subset \mathfrak{g}$ be an open convex symmetric neighborhood the origin such that $U_H^{-1}U_H \subset \exp(V \cap \mathfrak{h})$ and $\exp|_V$ is a diffeomorphism. Given a neighborhood $U_H \subset H$ of the identity in H , there exists a sequence $(h_n) \in G$ such that $H = \bigcup_n h_n U_H$. Choose $\epsilon > 0$ with $\exp([-2\epsilon, 2\epsilon]X) \subseteq \exp(V)$. Since the interval $[0, \epsilon]$ is uncountable, there exists some n for which $\{t \in [0, \epsilon] : e^{tX} \in h_n U_H\}$ is

uncountable. In particular, there are two such elements $t_1 < t_2$ and $u_1, u_2 \in U_H$ such that $e^{t_1 X} = h_n u_1 \neq e^{t_2 X} = h_n u_2$. Then

$$e^{(-t_1+t_2)X} = e^{-t_1 X} e^{t_2 X} = u_1^{-1} h_n^{-1} h_n u_2 \in U_H^{-1} U_H \subseteq \exp(V \cap \mathfrak{h}). \quad (12.6.2)$$

In particular, there is a nonzero $Y \in V \cap \mathfrak{h}$ with $e^Y = e^{(-t_1+t_2)X}$. Since Y and $(t_2 - t_1)X$ are contained in V , it follows that $X = \frac{1}{t_2 - t_1} Y \in \mathfrak{h}$. ■

Example 12.6.2 (Adjoint group). Let \mathfrak{g} be a finite-dimensional Lie algebra. Since $\text{ad} : \mathfrak{g} \rightarrow L(\mathfrak{g})$ is a homomorphism of Lie algebras, $\text{ad}(\mathfrak{g})$ is a Lie subalgebra of $L(\mathfrak{g}) = \text{Lie GL}(\mathfrak{g})$. According to the Analytic Subgroup Theorem, the subgroup of $\text{GL}(\mathfrak{g})$ generated by the $e^{\text{adx}}, X \in \mathfrak{g}$, is the unique connected Lie subgroup of $\text{GL}(\mathfrak{g})$ with Lie algebra $\text{ad}(\mathfrak{g})$. It is called the adjoint group Adg of the Lie algebra \mathfrak{g} .

By differentiating with respect to t and using the Jacobi identity, it is easy to check that the map $t \mapsto e^{-t\text{adx}} [e^{t\text{adx}} Y, e^{t\text{adx}} Z]$ is constant and equal to its value $[Y, Z]$ at $t = 0$, i.e.

$$[e^{\text{adx}} Y, e^{\text{adx}} Z] = e^{\text{adx}} [Y, Z] \text{ for all } X, Y, Z \in \mathfrak{g}; \quad (12.6.3)$$

and this shows that e^{adx} is an automorphism of \mathfrak{g} for any $X \in \mathfrak{g}$. It follows that Adg is a subgroup of the group $\text{Aut}(\mathfrak{g})$ of all automorphisms of \mathfrak{g} . The elements of Adg are therefore also called the *inner automorphisms* of \mathfrak{g} . It is easy to see from the definition that $\text{Aut}(\mathfrak{g})$ is an analytic submanifold, hence a closed Lie subgroup, of $\text{GL}(\mathfrak{g})$. Furthermore, the Lie algebra of $\text{Aut}(\mathfrak{g})$ is exactly the algebra of derivations

$$\text{Der}(\mathfrak{g}) = \{\phi \in L(\mathfrak{g}) \mid \forall X, Y \in \mathfrak{g} \ \phi([X, Y]) = [\phi(X), Y] + [X, \phi(Y)]\}. \quad (12.6.4)$$

Notice that $\text{adg} \subset \text{Der}(\mathfrak{g})$. In general Adg need not be closed in $\text{Aut}(\mathfrak{g})$, hence in $\text{GL}(\mathfrak{g})$, whereas on the other hand $\text{Aut}(\mathfrak{g})$ need not be connected.

Further it should be observed that if G is a Lie group with Lie algebra \mathfrak{g} , then Ad is a homomorphism of Lie groups $G \rightarrow \text{GL}(V)$ with $(\text{Ad})_{*e} = \text{ad}$, so it follows that Ad actually maps G_0 homomorphically onto Adg . The point is that nonisomorphic connected Lie groups with isomorphic Lie algebras do exist; the result above shows that their adjoint groups then are isomorphic and can be defined intrinsically in terms of the Lie algebras.

In Theorem 11.5.8(b) we have seen that $g e^X g^{-1} = e^{\text{Ad}_g X}$ for any $g \in G, X \in \mathfrak{g}$. So if $g \in \ker \text{Ad}$, then $ghg^{-1} = h$ for all $h \in \exp(\mathfrak{g})$, hence also for all $h \in G_0$ (the subgroup generated by $\exp(\mathfrak{g})$). Conversely, this implies that $g \in \ker \text{Ad}$. In particular, $\ker \text{Ad} \cap G_0 = Z(G_0)$, the center of G_0 . From Example 13.2.2(1) below we get that $Z(G_0)$ is a closed Lie subgroup of G_0 and that the surjection $\text{Ad} : G_0 \rightarrow \text{Adg}$ induces an isomorphism of Lie groups $G_0/Z(G_0) \cong \text{Adg}$.

Now we can provide an alternative proof of Theorem 11.3.10 which stated that every Lie group homomorphism with a simply connected domain is uniquely determined by its differential at the identity.

Second proof of Theorem 11.3.10 [33, Cor. 1.10.5]. The condition that ϕ is a homomorphism of Lie algebras just means that its graph is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$, which is regarded as the Lie algebra of $G \times H$. According to the Analytic Subgroup Theorem there is a connected Lie subgroup $F \subset G \times H$ having the graph of ϕ as its Lie algebra.

Let $\pi_1 : F \rightarrow G$ be the homomorphism defined as the composition of the inclusion and the projection $F \hookrightarrow G \times H \xrightarrow{\text{pr}_1} G$. Its differential $T_e\pi_1$ is equal to the projection of $\text{graph}(\phi) \subset \mathfrak{g} \times \mathfrak{h}$ onto the first variable, which is invertible. Since the differentials at all points of a Lie group are related to each other by an isomorphism, we get that $T_f\pi_1$ is invertible for all $f \in F$. So $\pi_1(F)$ is an open subgroup of G , and G is connected, thus $\pi_1(F) = G$ by Proposition 6.3.6. Since F is also connected, π_1 is a covering homomorphism of Lie groups $F \rightarrow G$, and thus by assumption π_1 is an isomorphism. By the InMT, π_1^{-1} is analytic, hence a homomorphism of Lie groups $G \rightarrow F$. Writing π_2 for the composition $F \hookrightarrow G \times H \xrightarrow{\text{pr}_2} H$, we get that $\Phi = \pi_2 \circ \pi_1^{-1}$ is a homomorphism of Lie groups $G \rightarrow H$, having F as its graph, and $T_e\Phi = \phi$ because $T_{(e,e)}F = \text{graph}(\phi)$.

Finally the uniqueness of Φ follows from the observation that if Φ is a homomorphism as desired, then its graph is a connected Lie subgroup of $G \times H$ having $\text{graph}(\phi)$ as its tangent space at (e, e) , so it follows from the uniqueness in the Analytic Subgroup Theorem that $\text{graph}(\Phi) = F$. ■

So far to prove the uniqueness of analytic structures on Lie groups we only had to assume C^2 a priori differentiability (enough to define the map ad , which is a second derivative of the conjugation map Ad). Now we can weaken analyticity assumptions in Lie groups not only to C^2 differentiability, but even to assumptions of purely topological nature, like continuity, still getting the same end results. The following is a typical example of this principle.

Theorem 12.6.3 ([33, Thm 1.10.6]). *Let H_0 be a closed subset of G that at the same time is a “local subgroup” of G in the sense that $e \in H_0$ and there is a symmetric neighborhood V of $e \in G$ such that $g^{-1} \in H_0$ and $gh \in H_0$ for all $g, h \in V \cap H_0$. Then there is a Lie subalgebra $\mathfrak{h} < \mathfrak{g} = \text{Lie } G$ such that $\exp(U)$ is a neighborhood of e in H_0 for each sufficiently small neighborhood U of the origin in \mathfrak{h} . In particular, H_0 is a neighborhood of e in the connected Lie subgroup H of G whose Lie algebra is \mathfrak{h} .*

Proof. Since the statement is totally local, we may restrict ourselves to a logarithmic chart and make use of a norm $|\cdot|$ on \mathfrak{g} .

Let \mathfrak{h} be the set of $X \in \mathfrak{g}$ for which there exist sequences $(h_n) \in H_0$ and $t_n \in [0, \infty)$ such that

$$\lim_{n \rightarrow \infty} h_n = 1 \text{ in } H_0 \quad (12.6.5)$$

and

$$\lim_{n \rightarrow \infty} t_n \ln h_n = X \text{ in } \mathfrak{g}. \quad (12.6.6)$$

We first show that there is a neighborhood $U \subset \mathfrak{h}$ of $0 \in \mathfrak{h}$ such that $\exp(U) \subset H_0$. By continuity of \exp at 0, there is an $\delta > 0$ such that $e^X \in V$ whenever $|X| < \delta$. Choose $0 < \epsilon < \delta$ and select $X \in U = \{X \in \mathfrak{h} \mid |X| \leq \epsilon\}$. Let h_n, t_n be as above. Since $\lim_n \ln h_n = 0$, we get $(t_n - \lfloor t_n \rfloor) \ln h_n \rightarrow 0$; so $\lfloor t_n \rfloor \ln h_n \rightarrow X$. If $0 \leq k \leq \lfloor t_n \rfloor - 1$, then:

$$\begin{aligned} |k \ln h_n| &= \frac{k}{\lfloor t_n \rfloor} |\lfloor t_n \rfloor \ln h_n| \leq \frac{k}{\lfloor t_n \rfloor} (|X| + |\lfloor t_n \rfloor \ln h_n - X|) \leq \\ &\leq \epsilon + |\lfloor t_n \rfloor \ln h_n - X| < \delta, \end{aligned} \quad (12.6.7)$$

for sufficiently large n . So $h_n^k = \exp(k \ln h_n) \in V$, for $0 \leq k \leq \lfloor t_n \rfloor - 1$; but now $h_n \in H_0 \cap V$ yields that $h_n^{\lfloor t_n \rfloor} \in H_0$ because of the assumptions. Using the continuity of \exp , we get that

$$\exp X = \lim_{n \rightarrow \infty} \exp(\lfloor t_n \rfloor \ln h_n) = \lim_{n \rightarrow \infty} h_n^{\lfloor t_n \rfloor}, \quad (12.6.8)$$

and the limit belongs to H_0 since H_0 is closed. Thus proves $\exp(U) \subset H_0$.

Next we show that \mathfrak{h} is a linear subspace of \mathfrak{g} . Obviously $tX \in \mathfrak{h}$ for any $X \in \mathfrak{h}$ and $t \geq 0$. Also $-X \in \mathfrak{h}$ because $V \cap H_0$ is symmetric. If now $X, Y \in \mathfrak{h}$ then by the result just proved $e^{tX}, e^{tY} \in H_0 \cap V$ for all sufficiently small t ; hence $e^{tX}e^{tY} \in H_0$ for small t . On the other hand, from the first term of the BCH formula,

$$\lim_{t \downarrow 0} \frac{1}{t} \ln e^{tX}e^{tY} = X + Y, \quad (12.6.9)$$

from which we read off that $X + Y \in \mathfrak{h}$.

Thirdly we prove that for any neighborhood $U \subset \mathfrak{h}$ of the origin the set $\exp(U)$ contains a neighborhood of the identity in H_0 . For if not, then there exists a sequence $h_n \in H_0 \setminus \exp(U)$ such that $h_n \rightarrow e$. Choose a linear subspace \mathfrak{m} of \mathfrak{g} such that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. Then $(X, Y) \mapsto e^X e^Y$ is a diffeomorphism from a neighborhood of $(0, 0)$ in $\mathfrak{h} \times \mathfrak{m}$ to a neighborhood of $e \in G$. Hence there exist sequences $X_n \in \mathfrak{h}, Y_n \in \mathfrak{m}$, both converging to 0, such that $h_n = e^{X_n} e^{Y_n}$. Note that $Y_n \neq 0$ because $h_n \notin \exp(U)$. Since $h_n \in H_0$ and $e^{X_n} \in H_0$, it follows that $e^{Y_n} \in H_0$ for sufficiently large n , according to the assumption of the theorem. On the other hand, a subsequence of $Y_n/|Y_n|$ converges to some $Y \in \mathfrak{m}$ with $|Y| = 1$. By definition of \mathfrak{h} we conclude that $Y \in \mathfrak{h}$, a contradiction with $\mathfrak{h} \cap \mathfrak{m} = 0$.

Taking $U = U' \cap \mathfrak{h}$ with U' a sufficiently small open neighborhood of o in \mathfrak{g} such that $\exp|_{U'}$ is a diffeomorphism from U' onto an open neighborhood $V' \subset G$ of $e \in G$, one obtains that $V = \exp(U)$ is a real-analytic submanifold of G and $\exp|_U$ is a real-analytic diffeomorphism from U onto V . Because on the other hand V is an open neighborhood of e in H_0 , it follows that V is a real-analytic local Lie group, with Lie algebra \mathfrak{h} , which apparently is a Lie subalgebra of \mathfrak{g} . If H denotes the group generated by V , then H is the connected Lie subgroup of G with Lie algebra equal to \mathfrak{h} that was introduced in the Analytic Subgroup Theorem 12.6.1. ■

We can now prove a slightly relaxed version of the Closed Subgroup Theorem 12.4.2.

Corollary 12.6.4 (Closed Subgroup Theorem [33, Cor. 1.10.7]). *Let H be a subgroup of a Lie group G . Then TFAE:*

- (i) *For some $h \in H$ there is a closed neighborhood V of h in G such that $H \cap V$ is closed in G .*
- (ii) *H is a closed Lie subgroup of G . The Lie group structure of H is unique if we require that its topology as a Lie group coincide with the relative (subspace) topology induced from G .*

Proof. Suppose (i) is valid. Applying the homeomorphism L_h^{-1} , we see that we may assume $h = e$ in (i). Now let $g \in G$ be the limit of a sequence $g_n \in G$. Then

$$\lim_{n,m \rightarrow \infty} g_n g_m^{-1} = e \Rightarrow g_n g_m^{-1} \in V \text{ for sufficiently large } n, m. \quad (12.6.10)$$

It follows that

$$g g_m^{-1} = \lim_{n \rightarrow \infty} g_n g_m^{-1} \in H \cap V \text{ for sufficiently large } m. \quad (12.6.11)$$

This implies that $g \in H$ because $g_m \in H$ and H is a group. This shows that H is closed in G . From Theorem 12.6.3 we see that H is a Lie subgroup of G ; actually the proof of Theorem 12.6.3 becomes slightly simpler if we assume that H is a closed subgroup of G instead of only a closed local group. Note that in Theorem 12.6.3 the exponential map is a homeomorphism from U to a neighborhood of e in H with respect to the relative topology making the Lie group topology of H equal to the relative topology. ■

Remark 12.6.5. It is also known that H is a Lie subgroup of G if H is a subgroup of a Lie group G such that, for each $h \in H$, there is a continuous curve $\gamma : [0, 1] \rightarrow G$ such that $\gamma(0) = e$, $\gamma(1) = h$, and $\gamma(t) \in H$ for all $t \in [0, 1]$. Here the Lie group structure of H is unique if we require that H is connected with respect to its Lie group topology. The proof is by Yamabe (1950).

12.7 Yamabe's Theorem

The central result of this section is Yamabe's theorem, characterizing analytic subgroups of a Lie group by the purely topological property of being path-connected. Groups which are not path-connected can then be dealt with using the concept of initial (weakly embedded) subgroups.

Theorem 12.7.1 (Yamabe [40, Thm. 9.6.1]). *A subgroup H of a connected Lie group G is path-connected iff it is an analytic subgroup, i.e. it is of the form $\langle \exp_G \mathfrak{h} \rangle$ for the Lie subalgebra \mathfrak{h} of $\text{Lie } G$ determined by*

$$\mathfrak{h} = \{X \in \text{Lie } G : \exp_G(\mathbb{R} \cdot X) \subset H\}. \quad (12.7.1)$$

Proof. We will not present the proof of the strong version of the theorem here. Instead we will prove it under the additional assumption (following [17, Lem. 1.7.10]) that H is C^∞ -pathwise connected, meaning that any two points in H can be connected by a path inside H that is smooth as a path in G . Consider the set

$$\mathfrak{h} = \{\dot{h}(0) \in T_e G : h \in C^\infty(\mathbb{R}, G), h(\mathbb{R}) \subset H, h(0) = e\}. \quad (12.7.2)$$

It is easy to check that this is a Lie subalgebra of \mathfrak{g} . Let \tilde{H} be the corresponding connected Lie subgroup of G provided by the Analytic Subgroup Theorem 12.6.1. As shown in the proof of that theorem, \tilde{H} is the maximal integral submanifold through e of the distribution $D^{\mathfrak{h}}$ generated by \mathfrak{h} . Now, let $h(t)$ be a smooth curve in G such that $h(\mathbb{R}) \subset H$ and $h(0) = e$. Clearly,

$$L_{h(t)*}^{-1} \dot{h}(t) = \frac{d}{ds} \Big|_{s=0} h(t)^{-1} h(t+s) \in \mathfrak{h}. \quad (12.7.3)$$

Thus, by left invariance of $D^{\mathfrak{h}}$, $t \mapsto h(t)$ lies in \tilde{H} . Since, by assumption, every point in H is connected to e by a smooth curve, we conclude $H \subset \tilde{H}$.

To prove $\tilde{H} \subset H$, choose a basis $(v_1, \dots, v_i) \in \mathfrak{h}$ and a family of smooth curves $t \mapsto h_i(t)$ such that $h_i(\mathbb{R}) \subset H$, $h_i(0) = e$ and $\dot{h}_i(0) = v_i$. Consider the mapping

$$F : \mathbb{R}^n \rightarrow \mathbb{R}, \quad F(\mathbf{t}) := h_1(t_1) \cdots h_n(t_n). \quad (12.7.4)$$

Clearly, $F'(0)$ maps \mathbb{R}^n bijectively onto \mathfrak{h} . Thus, by the InMT, F is a local diffeomorphism mapping an open neighborhood of the origin in \mathbb{R}^n onto an open e -neighborhood in H . We conclude that H contains an open e -neighborhood in \tilde{H} , and, thus $\tilde{H} \subset H$. ■

The following theorem, based on Yamabe's Theorem, states that *every* (set-theoretic) subgroup of a Lie group is in fact weakly embedded, completely dismantling our definition of an (immersed) Lie subgroup.

Theorem 12.7.2 (Initial Subgroup Theorem [40, Thm. 9.6.13]). *Each subgroup H of a Lie group G carries a structure of a weakly embedded (initial) Lie subgroup for which the identity component H_0 coincides with the path-component of H with respect to the subspace topology and*

$$\mathfrak{h} = \{X \in \text{Lie } G : \exp_G(\mathbb{R} \cdot X) \subset H\}. \quad (12.7.5)$$

Proof. Let $H_a \subset H$ be the path-component of H , viewed as a topological subgroup of G . According to Yamabe's Theorem, H_a is an analytic subgroup with Lie algebra

$$\{\mathfrak{h} = \{X \in \mathfrak{g} : \exp(\mathbb{R} \cdot X) \subseteq H\}\}. \quad (12.7.6)$$

Let H_a^L denote the group H_a endowed with its intrinsic Lie group topology for which $\exp = \exp_G|_{\mathfrak{h}} : \mathfrak{h} \rightarrow H_a^L$ is a local diffeomorphism at 0. Then $H \subseteq \{g \in G : \text{Ad}_g \mathfrak{h} = \mathfrak{h}\}$ and $\text{Ad}_g \circ \exp = \exp \circ \text{Ad}_g|_{\mathfrak{h}}$ imply that for each $h \in H$ the conjugation Ad_h defines a smooth automorphism on H_a^L , so that H carries a Lie group structure for which H_a^L is an open subgroup (see [40, Thm. 9.4.4, Cor. 9.4.5]). Let H^L denote this Lie group. Now the inclusion map $i : H^L \rightarrow G$ is an immersion whose differential at e is the inclusion $\mathfrak{h} \hookrightarrow \mathfrak{g}$.

We claim that $i : H^L \rightarrow G$ is a weakly embedded Lie subgroup. In fact, let $f : M \rightarrow G$ be a smooth map from the smooth manifold M to G with $f(M) \subseteq H$. We need to show that f is smooth. It suffices to show smoothness at each point $m \in M$ and by replacing f with $f \cdot f(m)^{-1}$ we can assume $f(m) = e$.

Let $U \subset \mathfrak{h}$ be an open 0-neighborhood, $\mathfrak{m} \subseteq \mathfrak{g}$ be a vector space complement to \mathfrak{h} , and $V \subseteq \mathfrak{m}$ an open 0-neighborhood for which the map $\Phi : U \times V \rightarrow G$ given by $(X, Y) \mapsto e^X e^Y$ is a diffeomorphism. Then

$$H_a \cap (\exp_G U \exp_G V) = \bigsqcup_{y \in V, \exp_G Y \in H_a} \exp_G U \cdot e^Y, \quad (12.7.7)$$

where the union is disjoint because Φ is bijective and each set $\exp_G U \cdot e^Y$ contained in H_a also is an open subset of H^L . Since the topology of H_a^L is second countable, the set $\exp_G^{-1}(H_a) \cap V$ is countable. Every smooth path $\gamma : I \rightarrow \exp_G U \exp_G V$ is of the form $\gamma(t) = \exp_G \alpha(t) \exp_G \beta(t)$ with smooth arcs $\alpha : I \rightarrow U$ and $\beta : I \rightarrow V$, and for every smooth path contained in H_a , the countability of $\beta(I)$ implies that β is constant.

We conclude that if $W \subseteq M$ is an open connected neighborhood of m with $f(W) \subseteq \exp_G U \exp_G V$, then $f(W) \subseteq \exp_G U$. Then the map $\exp_G|_U^{-1} \circ f|_W : W \rightarrow \mathfrak{h}$ is smooth, so that the corresponding restriction in range

$$i^{-1} \circ f|_W = \exp_{H^L} \circ \exp_G|_U^{-1} \circ f|_W : W \rightarrow H^L \quad (12.7.8)$$

is also smooth. This proves that the map $i^{-1} \circ f : M \rightarrow H^L$ is smooth, and hence that $i : H^L \rightarrow G$ is a weakly embedded Lie subgroup of G . ■

The Analytic and Closed Subgroup Theorems demonstrate that the connection between Lie subalgebras and Lie subgroups is subtle. Given a Lie group G with Lie algebra \mathfrak{g} , Lie subalgebras of \mathfrak{g} generate analytic subgroups of G , via integral manifolds or the exponential function, but unless they are closed, these groups are Lie groups only if one changes the topology (that is, they are not embedded but merely weakly embedded). On the other hand, the Initial Subgroup Theorem shows that any subgroup H can be given a topology that makes it a Lie group whose connected component arises in this way. Thus in defining the concept of a Lie subgroup one has to decide whether one wants to include the topology into the structure or not. If not, any subgroup is a Lie subgroup and the concept is superfluous. If one makes topology part of the structure, only closed subgroups will be Lie subgroups. Of course, this does not mean that the nonclosed subgroups associated with Lie subgroups are not important. Path-connected subgroups often go, for instance, under the name analytic subgroups. Singling out this class of subgroups is, of course, motivated by Yamabe's Theorem. If one wants to characterize the Lie group structure of general subgroups, in view of the Initial Subgroup Theorem, the initial submanifold property is the crucial one.

In all of these results we still have the overarching Lie group G that needs to be at least C^2 -differentiable for our derivations to result in an analytic structure. A series of even more advanced results that we will not cover here circumvent even this requirement. For example, one can show that any topological group that is also a topological manifold has a unique smooth structure turning it into a Lie group. This is essentially Hilbert's Fifth Problem, which can also be restated as follows: every continuous function $f : V \times V \rightarrow U$ between domains $V \subseteq U \subseteq \mathbb{R}^n$ that satisfies the group associativity law must be smooth (up to continuous reparametrization). This is a result of Montgomery-Zippin, Gleason, and Yamabe, and a concise proof can be found on [Terry Tao's blog](#).

13 Group Actions on Manifolds

13.1 Homogeneous spaces

Recall that every subgroup $H < G$ defines an equivalence relation on G by $a \sim_H b$ iff $a^{-1}b \in H$. The equivalence classes are given by the left cosets $aH \subset G$, $a \in G$. Alternatively, there is another equivalence relation $a \sim_H b \Leftrightarrow ab^{-1} \in H$ whose equivalence classes are the right cosets Ha , $a \in G$. By definition of the quotient topology the natural projections

$$\pi_L : G \rightarrow G/H, \pi(a) = aH; \quad \pi_R : G \rightarrow H\backslash G, \pi(a) = Ha \quad (13.1.1)$$

are continuous. Left translations descend to left cosets and right translations descend to right cosets via

$$\hat{L}_a : G/H \rightarrow G/H, \hat{L}_a(bH) = abH; \quad \hat{R}_a : H\backslash G \rightarrow H\backslash G, \hat{R}_a(Hb) = Hab \quad (13.1.2)$$

and satisfy

$$\hat{L}_a \circ \pi_L = \pi_L \circ L_a, \quad \hat{R}_a \circ \pi_R = \pi_R \circ \hat{R}_a. \quad (13.1.3)$$

The two quotients coincide only in Abelian groups. The topological structure on the quotient can be quite unpleasant. For example if $G = \mathbb{T}^2$ and $H = \{(e^{it}, e^{i\alpha t}) : t \in \mathbb{R}\}$ with $\alpha \in \mathbb{R}$ irrational, the quotient is not even a Hausdorff space. If, however, H is closed, then the quotient turns out to be a smooth manifold. In this section we will only consider left cosets, although all results will also hold for right cosets.

Lemma 13.1.1 ([16, Lem. 5.7.1]). *Let G be a Lie group and $H \subset G$ a (topological) subgroup.*

1. *The quotient map $\pi : G \rightarrow G/H$ is open.*
2. *The induced mappings \hat{L}_a are homeomorphisms of G/H .*
3. *If H is closed, then G/H is Hausdorff.*

Proof. 1. Let $U \subset G$ be open. Then, by definition of quotient topology, $\pi(U)$ is open iff $\pi^{-1}(\pi(U))$ is open in G . The latter holds because $\pi^{-1}(\pi(U)) = \bigcup_{a \in H} R_a(U)$ and the R_a are homeomorphisms of G .

2. Since $\hat{L}_a \circ \hat{L}_{a^{-1}} = \text{id}_{G/H}$, it suffices to show that \hat{L}_a is open for all $a \in G$. Let $\hat{U} \subset G/H$ be open. We have $\hat{L}_a(\hat{U}) = \pi \circ L_a(\pi^{-1}(\hat{U}))$. Since π is continuous and since L_a and π are open, the assertion follows.
3. Consider the mapping $\varphi : G \times G \rightarrow G$ defined by $\varphi(a, b) = a^{-1}b$. We have $\pi(a) = \pi(b)$ iff $(a, b) \in \pi^{-1}(H)$. Now let $\hat{a}_1 \neq \hat{a}_2 \in G/H$ and choose $a_i \in G$ such that $\pi(a_i) = \hat{a}_i$. Then $(a_1, a_2) \notin \varphi^{-1}(H)$. Since $\varphi^{-1}(H)$ is closed, there exist neighborhoods V_i of a_i

in G such that $(V_1 \times V_2) \cap \varphi^{-1}(H) = \emptyset$. This implies $\pi(V_1) \cap \pi(V_2) = \emptyset$. Since π is open, $\pi(V_i)$ is a neighborhood of \hat{a}_i , proving the Hausdorff property. ■

Theorem 13.1.2 (Homogeneous space). *Let G be a Lie group and let $H \subset G$ be a closed subgroup. There exists a unique smooth structure on the quotient G/H such that the natural projection $\pi : G \rightarrow G/H$ is a smooth submersion.*

Proof. It suffices to prove existence, because uniqueness holds for any submersion. Choose a subspace $\mathfrak{m} \subset \mathfrak{g}$ complementary to \mathfrak{h} . The manifold structure of G/H will be modelled on this subspace.

According to Lemma 12.4.1, there exist open neighborhoods $W_{\mathfrak{h}}$ and $W_{\mathfrak{m}}$ of the origins in \mathfrak{h} and \mathfrak{m} , respectively, and an open neighborhood V of $e \in G$ such that the mapping

$$\varphi : W_{\mathfrak{m}} \times W_{\mathfrak{h}} \rightarrow V, \quad (X, Y) \mapsto e^X e^Y, \quad (13.1.4)$$

is a diffeomorphism. Since H is closed and hence embedded, $\exp(W_{\mathfrak{h}})$ is open with respect to the relative topology on H induced from G . As a consequence, $W_{\mathfrak{m}}$ may be shrunk so that $(\exp(W_{\mathfrak{m}}))^2 \cap H \subset \exp(W_{\mathfrak{h}})$. Consider the mapping $\pi \circ \exp : W_{\mathfrak{m}} \rightarrow G/H$. It is

- (a) injective: if $X, Y \in W_{\mathfrak{m}}$ satisfy $e^{-X} e^Y \in H$, then $e^{-X} e^Y = e^Z$ for some $Z \in W_{\mathfrak{h}}$, hence $\varphi(Y, 0) = \varphi(X, Z)$ and injectivity of φ implies $X = Y$ (and $Z = 0$);
- (b) open: if U is open in $W_{\mathfrak{m}}$, then $U \times W_{\mathfrak{h}}$ is open in $W_{\mathfrak{m}} \times W_{\mathfrak{h}}$. Since φ and π are open, $\pi \circ \exp(U) = \pi \circ \varphi(U \times W_{\mathfrak{h}})$ is open in G/H .

Thus $\pi \circ \exp$ maps $W_{\mathfrak{m}}$ homeomorphically onto the subset $\hat{U} = \pi \circ \exp(W_{\mathfrak{m}}) \subset G/H$ and hence induces a local chart $\hat{\kappa} = (\pi \circ \exp)^{-1} L \hat{U} \rightarrow \mathfrak{m}$ on G/H . Since the mappings \hat{L}_a are homemorphisms, the collection

$$\{(\hat{L}_a(\hat{U}), \hat{\kappa} \circ \hat{L}_{a^{-1}}) \mid a \in G\} \quad (13.1.5)$$

is an atlas on G/H . The transition maps can be expressed in terms of \exp, \exp^{-1} , left translations and the projection $\mathfrak{g} \rightarrow \mathfrak{m}$ defined by the vector space decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. It follows that the atlas so constructed defines a smooth (in fact analytic) structure on G/H . To see that π is a smooth submersion, due to $\pi(L_a(V)) = \hat{L}_a(\hat{U})$, it suffices to check that the restriction $\pi : L_a(V) \rightarrow \hat{L}_a(\hat{U})$ is a smooth submersion for all $a \in G$. The latter follows from the fact that the mapping $\hat{\kappa} \circ \hat{L}_{a^{-1}} \circ \pi \circ L_a \circ \varphi : W_{\mathfrak{m}} \times W_{\mathfrak{h}} \rightarrow W_{\mathfrak{m}}$ coincides with the natural projection of the direct product. ■

Remark 13.1.3. If H is compact, then $\pi : G \rightarrow G/H$ is proper and thus, by the Ehresmann Fibration Theorem 7.1.8, a locally trivial fibration. The assumption of compactness on G can

in fact be dropped, as we will see in Section 13.7.

Remark 13.1.4. According to Theorem 3.10.3, the submersion π admits local sections. In the case of G/H there is a distinguished class of local sections. For every subspace \mathfrak{m} complementary to \mathfrak{h} there exist an open neighborhood W of the origin in \mathfrak{m} such that $\exp(W) \subset G$ is an embedded submanifold, $\pi|_{\exp(W)}$ is injective and $\hat{U} = \pi \circ \exp(W)$ is open in G/H . Then $s = \pi^{-1} : \hat{U} \rightarrow \exp(W)$ is a local section of π , and so are $s_a = L_a \circ s \circ L_{a^{-1}} : \hat{L}_a(\hat{U}) \rightarrow G$ for all $a \in G$. This argument also shows that the natural isomorphism $\mathfrak{g} \cong T_e G$ descends to an isomorphism

$$\mathfrak{g}/\mathfrak{h} \cong T_{[e]}(G/H). \quad (13.1.6)$$

The existence of local sections of π implies that the equality $\hat{L}_a \circ \pi = \pi \circ L_a$ can be solved locally for \hat{L}_a . This entails

Corollary 13.1.5. *The mapping \hat{L}_a is a diffeomorphism of G/H for all $a \in G$.*

Therefore each quotient G/H by a closed subgroup inherits a smooth action of G . It is clear that this action is transitive and the stabilizer group of each point is isomorphic to H . Conversely, the stabilizer of any point of a manifold M under a transitive action of G is a closed subgroup of G , and the stabilizers of different points are conjugates of each other. Therefore there is a one-to-one correspondence between manifolds with transitive actions of G and conjugacy classes of closed subgroups of G . These pairs deserve a special name.

Definition 13.1.6 (Homogeneous space). A homogeneous space (also historically called a *Klein geometry*) is a manifold with a transitive action of a Lie group. Alternatively, it is a manifold obtained as a (right or left) quotient of a Lie group G by a closed subgroup H , together with the canonical (resp. left or right) G -action on it. We call it a *homogeneous space of type (G, H)* .

Example 13.1.7 (Principal homogenous spaces, Affine spaces). 1. Putting $H = \{e\}$, we get that the so called *principal homogeneous space* $G/\{e\}$ is isomorphic to G as a manifold, but doesn't have a group structure. This manifold can be defined as the manifold with a *free and transitive* G -action. However, for any two points $x, y \in G/\{e\}$ there is a unique $g \in G$ such that $y = g \circ x$. Thus elements of $G/\{e\}$ can be "divided" but not multiplied, and there is no identity. Principal homogeneous spaces are also called *G -torsors*.
 2. An *affine space* is nothing but the homogeneous space $A(V) = V/\{0\}$ for a vector space V viewed as an abelian Lie group. Its elements can be non-canonically identified with V : any two points $x, y \in A(V)$ differ from each other by a vector $v \in V$, i.e. " $y = x + v$ ", but there is no (natural) origin and no structure of a vector space. The affine space $A(V)$

is also the homogenous space of the affine group $\text{Aff}(V) = V \rtimes \text{GL}(V)$ corresponding to the stabilizer subgroup $\text{GL}(V)$.

3. Every connected manifold is a homogeneous space of its own diffeomorphism group. Indeed, $\text{Diff}(M)$ acts transitively on M thus $M \cong \text{Diff}(M)/H$, where H is the subgroup of diffeomorphisms fixing some pre-chosen basepoint $m \in M$. While curious, this fact is barely of any use because diffeomorphism groups are extremely complicated.
4. If $G \xrightarrow{\sigma} F$ is a faithful action and F' is diffeomorphic to F , then let O be an orbit in the space $\text{Diff}(F', F)$ under the natural G -action $(g \cdot \chi)(p) = g \cdot \chi(p)$, where $g \in G, \chi \in \text{Diff}(F', F)$. This action is free since σ is faithful, and it's transitive when restricted to O . Thus O is a principal homogeneous space for G . In the context of bundles we have called the choice of O a G -structure on a fiber, and O itself the set of G -frames for that fiber, cf. 7.3.1.

Another very important example of homogeneous spaces are Stiefel manifolds, arising from classical Lie groups.

Definition 13.1.8 (Stiefel manifolds). Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, or \mathbb{H} . The Stiefel manifold $V_k(\mathbb{K}^n)$ is defined to be the set of k -frames (a k -frame is a set of k linearly independent vectors) in \mathbb{K}^n which are orthonormal with respect to the standard scalar product. The actual manifold is obtained by identifying $V_k(\mathbb{K}^n)$ with a certain homogeneous space as follows. Any orthonormal k -frame in \mathbb{K}^n can be obtained from the first k elements of the standard basis in \mathbb{K}^n by a linear isometric transformation. This transformation is determined by the given frame only up to its action on the last $n - k$ elements of the standard basis. That is, two transformations produce the same k -frame iff they differ by prior application of a block matrix

$$\begin{bmatrix} I_k & 0 \\ 0 & a \end{bmatrix}, \quad (13.1.7)$$

where a is an isometry of \mathbb{K}^n . Hence, as sets,

$$V_k(\mathbb{R}^n) = O_n/O_{n-k}, \quad (13.1.8)$$

$$V_k(\mathbb{C}^n) = U_n/U_{n-k}, \quad (13.1.9)$$

$$V_k(\mathbb{H}^n) = Sp_n/Sp_{n-k}, \quad (13.1.10)$$

and these equalities define the smooth structures of Stiefel manifolds.

Remark 13.1.9. For $k < n$, a possible sign (if $\mathbb{K} = \mathbb{R}$) or phase (if $\mathbb{K} = \mathbb{C}$) in the determinant of the isometry producing the desired k -frame can be shifted to the irrelevant part which acts on the last $n - k$ standard basis vectors, therefore

$$V_k(\mathbb{R}^n) = SO_n/SO_{n-k}, \quad V_k(\mathbb{C}^n) = SU_n/SU_{n-k}. \quad (13.1.11)$$

An orthonormal 1-frame is just a unit vector, hence

$$V_1(\mathbb{K}^n) \cong \mathbb{S}^{dn-1}, \quad d = \dim_{\mathbb{R}}(\mathbb{K}). \quad (13.1.12)$$

Moreover, $V_2(\mathbb{K}^n)$ can be identified with the space of unit tangent vectors to $V_1(\mathbb{K}^n)$ since a vector v at the point $x \in \mathbb{S}^m$ (where $m = dn - 1$) is tangent to \mathbb{S}^m iff it is orthogonal to x .

Similarly, an orthonormal n -frame of \mathbb{K}^n can be identified with the columns of a matrix representing an isometry, hence

$$V_n(\mathbb{R}^n) = \mathrm{O}_n, \quad V_n(\mathbb{C}^n) = \mathrm{U}_n \quad V_n(\mathbb{H}^n) = \mathrm{Sp}_n. \quad (13.1.13)$$

Moreover, $V_{n-1}(\mathbb{R}^n) \cong \mathrm{SO}_n$ and $V_{n-1}(\mathbb{R}^n) \cong \mathrm{SU}_n$ since there is a unique way of extending an orthonormal $(n-1)$ -frame to a positively oriented orthonormal n -frame. These diffeomorphisms also follow by noticing that (13.1.11) become principal homogeneous spaces when $k = n, n-1$.

When $k < n$, we can use the description of $V_k(\mathbb{R}^n)$ as a coset space to induce a CW -structure from the one on SO_n (described in Section 6.5). The cells are the sets of cosets of the form $e^I \mathrm{SO}_{n-k} = e^{i_1} \cdots e^{i_m} \mathrm{SO}_{n-k}$ for $n > i_1 > \cdots > i_m \geq n-k$, together with the coset SO_{n-k} itself as a 0-cell of V_k . These sets of cosets are unions of cells of SO_n since SO_{n-k} consists of the cells $e^J = e^{j_1} \cdots e^{j_l}$ with $n-k > j_1 > \cdots > j_l$. This implies that $V_k(\mathbb{R}^n)$ is the disjoint union of its cells, and the boundary of each cell is contained in cells of lower dimension, so we do have a CW -structure.

Definition 13.1.10 (Infinite Stiefel spaces). The coordinate inclusion maps $\mathbb{K}^n \hookrightarrow \mathbb{K}^{n+1}$ induce corresponding inclusions $V_k(\mathbb{K}^n) \hookrightarrow V_k(\mathbb{K}^{n+1})$. By taking the direct limit, we get their unions

$$V_k(\mathbb{K}^\infty) = \bigcup_n V_k(\mathbb{K}^n). \quad (13.1.14)$$

Note that $V_k(\mathbb{K}^\infty)$ can be thought of as a closed subspace of $\mathbb{S}^\infty \times \cdots \times \mathbb{S}^\infty$ (k factors). It also has a CW -structure induced from the finite-dimensional Stiefel spaces.

13.2 Quotients by normal subgroups

Next we investigate the special case where H is a normal subgroup. This means that $aH = Ha$, and thus G/H inherits a group structure.

Proposition 13.2.1 (Quotient Lie group [16, Prop. 5.7.8]). *Let G be a Lie group with Lie algebra \mathfrak{g} and let $H \subset G$ be a closed normal subgroup with associated Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then:*

1. *G/H with the induced group structure is a Lie group.*

2. The natural projection $\pi : G \rightarrow G/H$ is a Lie group homomorphism.
3. We have $\ker(\pi_{*e}) = \mathfrak{h}$ and π_{*e} induces an isomorphism from $\mathfrak{g}/\mathfrak{h}$ onto the Lie algebra of G/H .

Proof. 1. Since π is a group homomorphism, the multiplication mappings m of G and \hat{m} of G/H satisfy $\hat{m} \circ (\pi \times \pi) = \pi \circ m$. Using pairs of local sections of π , this equality can be locally solved for $\hat{\pi}$. Hence \hat{m} is smooth and G/H is a Lie group.

2. This holds by construction.
3. Since $\ker \pi = H$, and π is a Lie group homomorphism, we know that $\ker \pi_{*e} = \mathfrak{h}$ from Example 12.4.4 and the assertion follows from the homomorphism theorem for algebras. ■

- Example 13.2.2.**
1. The kernel of any Lie group homomorphism $\Phi : G \rightarrow H$ is a closed subgroup and hence $G/\ker \Phi$ is a Lie group with Lie algebra $\mathfrak{g}/\ker(\Phi_{*e})$.
 2. The identity component G_0 is a normal subgroup: Ad_g is a homeomorphism of G for any G , hence $\text{Ad}_g(G_0)$ is a connected component, but it also contains $e \in G$, hence $\text{Ad}_g(G_0) = G_0$. The cosets coincide with the connected components of G , hence G/G_0 is discrete.

Corollary 13.2.3 (Classification of abelian Lie groups). *Every connected abelian real Lie group, for some $n, k \geq 0$, is isomorphic to*

$$\mathbb{R}^n \times \mathbb{T}^k. \quad (13.2.1)$$

Proof. When G is abelian, the exponential map is actually a homomorphism of abelian groups (since $e^{X+Y} = e^X e^Y$ in this case). Being a homomorphism, it has constant rank (by Theorem 6.6.6) equal to the dimension of the group, therefore it is a submersion and hence open. Since any neighborhood of the identity generates the entire connected Lie group, any open homomorphism is surjective. Further, the fact that $\exp : \mathfrak{g} \rightarrow G$ is a local homeomorphism means that $\ker \exp$ is a discrete subgroup of \mathfrak{g} . By Corollary 6.2.10, this subgroup is a lattice Λ generated by a set of linearly independent vectors $x_1, \dots, x_k \in \mathfrak{g}$. Thus

$$G \cong \mathfrak{g}/\Lambda \cong \mathbb{R}^n \times \mathbb{T}^k, \quad (13.2.2)$$

where $n = \dim G - k$. ■

Proposition 13.2.4 (Quotient homomorphism [16, Prop. 5.7.10]). *Let G_i be Lie groups with Lie algebras \mathfrak{g}_i and let $H_i \trianglelefteq G_i$ be closed normal subgroups with associated Lie algebras $\mathfrak{h}_i \trianglelefteq \mathfrak{g}_i$, $i = 1, 2$. Let $\varphi : G_1 \rightarrow G_2$ be a Lie group homomorphism such that $\varphi(H_1) \subset H_2$. Then:*

1. There exists a unique mapping $\hat{\varphi} : G_1/H_1 \rightarrow G_2/H_2$ such that for the natural

projections $\pi_i : G_i \rightarrow G_i/H_i$, $i = 1, 2$, there holds $\hat{\varphi} \circ \pi_1 = \pi_2 \circ \varphi$. This mapping is a Lie group homomorphism.

2. There holds $\varphi_{*e}(\mathfrak{h}_1) < \mathfrak{h}_2$ and $\hat{\varphi}_{*e}$ is given by the homomorphism $\mathfrak{g}_1/\mathfrak{h}_1 \rightarrow \mathfrak{g}_2/\mathfrak{h}_2$ induced by φ_{*e} on passing to the quotients.

Proof. 1. Due to $\varphi(H_1) \subset H_2$, the map $\hat{\varphi}$ is well-defined. It is obviously unique. A brief calculation, using that π_i and φ are group homomorphisms, shows that $\hat{\varphi}$ is a homomorphism. Since π_1 admits local sections, $\hat{\varphi}$ is smooth.

2. By naturality of differentials (chain rule), we have $\hat{\varphi}_{*e} \circ (\pi_1)_{*e} = (\pi_2)_{*e} \circ (\varphi_{*e})$. Since under the identification of the Lie algebras of G_i/H_i with the quotients $\mathfrak{g}/\mathfrak{h}_i$, $(\pi_i)_{*e}$ corresponds to the natural projection $\mathfrak{g}_i \rightarrow \mathfrak{g}_i/\mathfrak{h}_i$, this implies the assertion. ■

This leads us to the Lie group analogue of the homomorphism theorem.

Proposition 13.2.5 (Homomorphism theorem [16, Prop. 5.7.11]). *Let $\varphi : G \rightarrow H$ be a Lie group homomorphism. Then $\text{im } \varphi$ is a Lie subgroup of H and the induced mapping $\hat{\varphi} : G/\ker \varphi \rightarrow \text{im } \varphi$ is a Lie group isomorphism.*

Proof. By combining Propositions 13.2.1 and 13.2.4, we have that $G/\ker \varphi$ is a Lie group and the induced mapping $\hat{\varphi} : G/\ker \varphi \rightarrow H$ is a Lie group homomorphism. By construction, $\hat{\varphi}$ is injective and hence an immersion by Corollary 6.1.9. Thus $\hat{\varphi} : G/\ker \varphi \rightarrow H$ defines a Lie subgroup of H and so is its image $\text{im } \varphi$ with respect to the smooth structure transported by $\hat{\varphi}$. ■

Example 13.2.6. 1. The exponential function $t \mapsto e^{2\pi i t}$ has kernel \mathbb{Z} and image U_1 . The induced mapping yields a Lie group isomorphism from \mathbb{R}/\mathbb{Z} onto U_1 . This extends to an isomorphism from $\mathbb{R}^n/\mathbb{Z}^n$ onto \mathbb{T}^n for any n .
 2. The two-fold covering map $SU_2 \rightarrow SO_3(\mathbb{R})$ of Example 6.1.10 is surjective and has the center \mathbb{Z}_2 of SU_2 as its kernel, hence the induced mapping yields a Lie group isomorphism $SU_2/\mathbb{Z}_2 \cong SO_3(\mathbb{R})$. Similarly, Example 6.4.5 yields $SL_2(\mathbb{C})/(\mathbb{C}^* I) \cong SO_{1,3}(\mathbb{R})$.

Exercise 13.1. Show that any normal connected subgroup H of a simply connected Lie group G is necessarily closed in G . *Hint:* let \mathfrak{g} be the Lie algebra of G and $\mathfrak{h} \trianglelefteq \mathfrak{g}$ the ideal corresponding to G . If B is a connected Lie group whose Lie algebra is isomorphic to $\mathfrak{g}/\mathfrak{h}$, prove that there is a continuous homomorphism $G \rightarrow B$ having H for the identity component of its kernel.

Exercise 13.2. Let H be a OPS of a Lie group G . Show that the closure of H in G is compact and thus a torus, if H is not closed in G (if it is closed, then it's obviously diffeomorphic either to \mathbb{R} or \mathbb{S}^1). *Hint:* use Corollary 12.6.4 to reduce to the case that G is connected and

abelian.

Exercise 13.3. Let G be a compact connected complex-analytic Lie group of dimension n . Prove that G is abelian and isomorphic to \mathbb{C}^n/Γ , where Γ is an integer lattice in \mathbb{C}^n with $2n$ generators (over \mathbb{R}). *Hint:* the adjoint representation $G \rightarrow \text{Ad}g$ is trivial since complex-analytic functions on a compact connected manifold are constant.

13.3 Lie group actions II

We begin by recalling a few definitions related to actions of Lie groups on manifolds (see Section 6.6).

Remark 13.3.1. If M is endowed with additional structure, one may require the diffeomorphisms Φ_g to respect this structure. This way one obtains, for example, the notion of Riemannian Lie group actions (by isometries), symplectic actions (by symplectomorphisms), and vector bundle actions (by vector bundle automorphisms).

Definition 13.3.2. Let (M, G, Φ) be a manifold M with a smooth action Φ of a Lie group G .

1. Two points $m_1, m_2 \in M$ are called conjugate under Φ if $m_2 = \Phi_g(m_1)$ for some $g \in G$. Conjugacy is an equivalence relation and its equivalence classes are called orbits of G under Φ . The set of equivalence classes, equipped with the quotient topology, is called the orbit space of (M, G, Φ) and is denoted by M/G (sometimes left and right actions are distinguished so that the quotient by a left action is denoted $G\backslash M$ and by a right action M/G).
2. For $m \in M$, the induced mapping $\Phi^m : G \rightarrow M$ is called the orbit map of m .
3. For $m \in M$, the subgroup $G_m = \{g \in G \mid \Phi_g = \text{id}_M\}$ of G is called the stabilizer, or the isotropy group, of m .
4. The kernel of Φ is the subgroup $\ker \Phi = \{g \in G \mid \Phi_g = \text{id}_M\}$ of G .

Assuming that Φ is a left action, the orbit map satisfies $\Phi^m(gh) = \Phi^m(g)$ iff $h \in G_m$, so it induces an injective map

$$\hat{\Phi}^m : G/G_m \rightarrow M, \quad \hat{\Phi}^m(gG_m) = \Phi^m(g), \quad (13.3.1)$$

where G/G_m denotes the homogeneous space of left cosets. Since the natural projection $G \rightarrow G/G_m$ is a surjective submersion, $\hat{\Phi}^m$ is smooth. Moreover, it is equivariant with respect to the action of G on G/G_m by the induced left translations \hat{L}_g and its image coincides with the orbit of m . Similarly, in the case of a right action, $\hat{\Phi}^m$ is defined on the homogeneous space of right cosets and is equivariant with respect to the induced right translations \hat{R}_g . Thus, for every m , $\hat{\Phi}^m$ yields an equivariant bijection between G/G_m and the orbit through m . We will use this bijection to induce a smooth structure on $G \cdot m$.

Some of the statements of the following Proposition should already be familiar.

Proposition 13.3.3 ([16, Prop. 6.1.5]). *Let (M, G, Φ) be a Lie group action.*

1. *The action $\Phi : G \times M \rightarrow M$ is an open mapping.*
2. *The natural projection $\pi : M \rightarrow M/G$ is open.*
3. *M/G is locally compact and second countable.*
4. *G_m is a closed subgroup of G for all $m \in M$. The mapping $g \mapsto (\Phi_g)_{*m}$ defines a representation of G_m on $T_m M$, called the isotropy representation at m .*
5. *The kernel of Φ is a closed subgroup of G . The mapping*

$$\hat{\Phi} : G/\ker \Phi \times M \rightarrow M, \quad \hat{\Phi} \circ (\rho \times \text{id}_M) = \phi, \quad (13.3.2)$$

where $\rho : G \rightarrow G/\ker \Phi$ denotes the natural projection, is an action of $G/\ker \Phi$ on M and (id_M, ρ) is a morphism of Lie group actions.

Proof. 1. This follows from $\Phi(V \times U) = \bigcup_{g \in V} \Phi_g(U)$ and the fact that the maps Φ_g are diffeomorphisms.

2. Let $U \subset M$ be open. Then, $\Phi_g(U)$ is open for all $g \in G$ and hence $\pi^{-1}(\pi(U)) = \bigcup_{g \in G} \Phi_g(U)$ is open. By construction of the quotient topology, then, $\pi(U)$ is open.
3. Since M is locally compact, every $m \in M$ has a compact neighborhood U . Since π is open, $\pi(U)$ is a neighborhood of $\pi(m)$ in M/G . Since the image of a compact space under a continuous mapping is compact, $\pi(U)$ is compact. Hence, M/G is locally compact. Second countability holds by definition of the quotient topology.
4. For every $m \in M$, the stabilizer G_m is the preimage of m under the orbit map Φ^m . The rest follows from the general facts that $(\Phi_g)_* : TM \rightarrow TM$ is also a Lie group action and that restricting to an initial invariant submanifold (such as $T_m M$ in this case) also produces a smooth action.
5. Being the kernel of a group (anti-)homomorphism, $\ker \Phi$ is normal. Being the intersection of the stabilizers of all points of M , it is also closed. Hence, by Proposition 13.2.1, $G/\ker \Phi$ is a Lie group. Since $\Phi_g = \text{id}_M$ for $g \in \ker \Phi$, $\hat{\Phi}$ is well-defined. Since ρ is a submersion, $\hat{\Phi}$ is smooth by the universal property of submersions (or by an argument via smooth local sections of ρ). The rest is obvious.

■

Next, we observe

$$G_{\Phi_g(m)} = gG_mg^{-1} \text{ (left action)}, \quad G_{\Phi_g(m)} = g^{-1}G_mg \text{ (right action)}. \quad (13.3.3)$$

That is, the stabilizers of any two points lying on the same orbit are conjugate.

Definition 13.3.4 (Orbit type). *Let (M, G, Φ) be a Lie group action. The type of an orbit O is defined to be the conjugacy class of subgroups of G given by $\{G_m : m \in O\}$. The set of*

conjugacy classes of subgroups of G which appear as types of orbits under Φ is called the set of orbit types of (M, G, Φ) . For a given orbit type $[H]$, the subset of M of orbit type $[H]$ will be denoted by $M_{[H]}$.

Definition 13.3.5. For a closed subgroup $H \subset G$, define

$$M_H = \{m \in M : H = G_m\}, \quad (13.3.4)$$

$$M^H = \{m \in M : H \subset G_m\}. \quad (13.3.5)$$

M_H and M^H are called the subset of isotropy type H and the subset of fixed points under H , respectively. By construction,

$$M_{[H]} = G \cdot M_H. \quad (13.3.6)$$

Moreover, if H is compact, we have

$$M_H = M^H \cap M_{[H]}. \quad (13.3.7)$$

We also reiterate definitions of the main algebraic properties of group actions.

Definition 13.3.6. A Lie group action Φ is called

1. effective if $\ker \Phi = \{e\}$,
2. free if all stabilizers are trivial: $\forall g \in G \quad G_m = \{e\}$.
3. transitive if it has a single orbit.

By the above Proposition, by passing to the induced action of $G/\ker \Phi$, every action can be replaced by an effective action which has the same orbits. Moreover, it is obvious that free actions and transitive actions have just one single orbit type.

Example 13.3.7. 1. The actions of G on itself by left and right translations are effective, free and transitive. Their restrictions to a proper Lie subgroup $H < G$ are still effective and free but no longer transitive. The action of G on itself by inner automorphisms (conjugations **Ad**) is neither free nor transitive, because the identity is invariant and the kernel is given by the center $Z(G)$. Hence, the adjoint action is effective iff the center is trivial. Its orbits are the conjugacy classes of elements of G . The stabilizer of $g \in G$ is the centralizer $C(g)$ of $g \in G$. Therefore, the orbit types are given by the conjugacy classes of those subgroups of G which are centralizers of elements of G . The orbit space is usually referred to as the adjoint quotient of G .

2. The action of G on the homogeneous space G/H , where $H \trianglelefteq G$ is a nontrivial closed subgroup, by the induced left translations \hat{L}_g is transitive but not free. The stabilizer of a coset gH , $g \in G$, is given by $G_{gH} = gHg^{-1}$. Consequently, the kernel of the action is given by $\bigcap_{g \in G} gHg^{-1}$. It lies between $Z(G) \cap H$ and H itself, the latter being the case when H is normal.

3. For the action of $G = \mathbb{R}$ on M defined by the flow F^X of a complete vector field $X \in \mathfrak{X}(M)$, the orbit map of a point $m \in M$ is given by the maximal integral curve F_m^X . The stabilizer is $G_m = \mathbb{R}$ in the case $X_m = 0$, $G_m = \mathbb{Z}T$ if the integral curve is periodic with period T , and $G_m = \{0\}$ otherwise. Since G is abelian, the orbit types correspond bijectively to the subgroups which appear as stabilizers.

13.4 Killing vector fields

Throughout this section, let (M, G, Φ) be a Lie group action and let \mathfrak{g} be the Lie algebra of G . For $A \in \mathfrak{g}$, we have

$$\Phi_{e^{tA}} \circ \Phi_{e^{sA}} = \Phi_{e^{(t+s)A}}, \quad (13.4.1)$$

hence the assignment $(m, t) \mapsto \Phi_{e^{tA}}(m)$ defines a OPS of diffeomorphisms of M , that is, a complete flow.

Definition 13.4.1 (Killing vector field). The vector field on M defined by the flow $\Phi_{e^{tA}}$ is called the Killing vector field generated by $A \in \mathfrak{g}$. It will be denoted by A_* . Its value at $m \in M$ may be expressed in several equivalent ways:

$$(A_*)_m = \frac{d}{dt} \Big|_{t=0} \Phi_{e^{tA}}(m) = \frac{d}{dt} \Big|_{t=0} \Phi^m(e^{tA}) = \Phi^m * e \cdot A. \quad (13.4.2)$$

Proposition 13.4.2 ([16, Prop. 6.2.2]). 1. For every $g \in G$ and $A \in \mathfrak{g}$, there holds $\Phi_{g*}A_* = (\text{Ad}_g A)_*$ in case Φ is a left action and $\Phi_{g*}A_* = (\text{Ad}_g^{-1} A)_*$ in case Φ is a right action.

2. The mapping $\mathfrak{g} \rightarrow \mathfrak{X}(M)$ given by $A \mapsto A_*$ is an anti-homomorphism of Lie algebras if Φ is a left action and a homomorphism if it is a right action. The kernel of this map coincides with the Lie algebra of $\ker \Phi \subset G$. In particular, if Φ is effective, the map $A \mapsto A_*$ is injective.
3. For $m \in M$, the Lie algebra of G_m is given by $\mathfrak{g}_m = \{A \in \mathfrak{g} : (A_*)_m = 0\}$. In particular, if Φ is free, then A_* cannot have critical points (zeros).

Proof. 1. By naturality of flows with respect to diffeomorphisms (Corollary 9.1.9), the flow of $\Phi_{g*}A_*$ is $\Phi_g \circ \Phi_{e^{tA}} \circ \Phi_{g^{-1}}$, which is equal to $\Phi_{e^{t\text{Ad}_g A}}$ if Φ is a left action and $\Phi_{e^{t\text{Ad}_g^{-1} A}}$ if Φ is a right action.

2. We give the proof for a left action. Linearity is obvious. Next we have

$$[A_*, B_*]_m = (\mathcal{L}_{A_*} B_*)_m = \frac{d}{dt} \Big|_{t=0} (\Phi_{e^{-tA_*}} B_*)_m = \Phi_{*e}^m \left(\frac{d}{dt} \Big|_{t=0} (\text{Ad}_{e^{-tA_*}} B) \right). \quad (13.4.3)$$

Now, the definition of the differential and the fact that $\text{Ad}_* = \text{ad}$ yield $[A_*, B_*] = -[A, B]_*$. Furthermore, A belongs to the Lie algebra of $\ker \phi$ iff $\Phi_{e^{tA}} = \text{id}_M$ for all t . Since $\Phi_{e^{tA}}$ is the flow of A_* , this is equivalent to $A_* = 0$.

3. Let $A \in \mathfrak{g}$. By the Closed Subgroup Theorem 12.4.2, A belongs to the Lie algebra \mathfrak{g}_m of G_m iff $\Phi_{e^{tA}}(m) = m$ for all t , that is, iff m is an equilibrium of A_* : $(A_*)_m = 0$.

■

Remark 13.4.3. According to the above Proposition, m is an equilibrium of A_* for all $A \in \mathfrak{g}_m$. Thus, the representation of \mathfrak{g}_m on $T_m M$ induced by the isotropy representation of G_m is given by

$$\mathfrak{g}_m \rightarrow \text{End}(T_m M), \quad A \mapsto \left. \frac{d}{dt} \right|_{t=0} (\Phi_{e^{tA}})_{*m} = \text{Hess}_m(A_*), \quad (13.4.4)$$

i.e. by the linearized flow (the Hessian) of the Killing vector field. It is referred to as the isotropy representation of \mathfrak{g}_m .

Proposition 13.4.4 (Transformation properties). *1. Let (M_i, G_i, Φ^i) , $i = 1, 2$, be Lie group actions and let $\varphi : M_1 \rightarrow M_2$, $\rho : G_1 \rightarrow G_2$ define a morphism (i.e. a commuting square $\Phi^2 \circ (\rho \times \varphi) = \varphi \circ \Phi^1$). Then the respective Killing vector fields of A on M_1 and of $\rho_* A$ on M_2 are φ -related.*

2. Let (M_i, Φ^i) , $i = 1, 2$, be G -manifolds and let $\varphi : M_1 \rightarrow M_2$ be equivariant. For every $A \in \mathfrak{g}$, the Killing vector fields $A_^{M_i}$ on M_i generated by A are φ -related.*

If φ is a diffeomorphism, points 1 and 2 together yield, respectively,

$$(\rho_{*e} A)_* = \varphi_* A_*, \quad A_*^{M_2} = \varphi_* A_*^{M_1}. \quad (13.4.5)$$

Proof. By naturality of the exponential map, for $m \in M_1$ we have

$$(\varphi_{*e} \circ A_*)(m) = \left. \frac{d}{dt} \right|_{t=0} \varphi \circ \Phi_{\exp_1(tA)}^1(m) \quad (13.4.6)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \Phi_{\rho(\exp_1(tA))}^2 \circ \varphi(m) \quad (13.4.7)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp_2(t\rho_{*e} A)}^2 \circ \varphi(m). \quad (13.4.8)$$

This yields the first assertion. The second one follows by letting $G_1 = G_2 = G$ and $\rho = \text{id}_G$. ■

Example 13.4.5. 1. For the left action of G associated with a representation $\rho : G \rightarrow \mathrm{GL}(V)$ on a finite-dimensional vector space V , the Killing vector field A_* is the linear vector field on V corresponding to the endomorphism $\rho_{*e}A$. In particular, for the adjoint and coadjoint actions we obtain

$$A_*^{\mathrm{Ad}} = \mathrm{ad}_A, \quad A_*^{\mathrm{Ad}^*} = \mathrm{ad}_A^*, \quad A \in \mathfrak{g}. \quad (13.4.9)$$

2. For the action of G on itself by right translations we get $A_* = A_L$. That is, the homomorphism $\mathfrak{g} \rightarrow \mathfrak{X}(G)$ of Proposition 13.4.2(2) coincides with the natural isomorphism between the Lie algebra and the left-invariant vector fields. The statement for left translations is analogous. For the adjoint action Ad , $A_* = A_R - A_L$, which agrees with the fact that $A \mapsto A_*$ should be an anti-homomorphism.
3. Let $H \subset G$ be a closed subgroup and let $\pi : G \rightarrow G/H$ denote the natural projection. For the action of G on the homogeneous space G/H by induced left translations, Proposition 13.4.4(2) and the previous result yield $(A_*)_{\pi(g)} = \pi_{*e} \circ (R_g)_{*e}A$ for all $g \in G$ and $A \in \mathfrak{g}$.
4. For the action of $G = \mathbb{R}$ on M defined by the flow of a complete vector field X , under the identification of the Lie algebra of \mathbb{R} with \mathbb{R} itself, the homomorphism $\mathfrak{g} \rightarrow \mathfrak{X}(M)$ is given by $s \mapsto s_* = sX$.

By analogy with Lie group actions and their morphisms one defines

Definition 13.4.6 (Lie algebra action). An action of a Lie algebra \mathfrak{g} on a smooth manifold M is a mapping $\phi : \mathfrak{g} \rightarrow \mathfrak{X}(M)$ which is either an anti-homomorphism (“left action”) or a homomorphism (“right action”) of Lie algebras. In the first case. The triple (M, \mathfrak{g}, ϕ) is referred to as a Lie algebra action and the pair (M, ϕ) as a \mathfrak{g} -manifold.

A morphism of Lie algebra actions $(M_1, \mathfrak{g}_1, \phi_1)$ and $(M_2, \mathfrak{g}_2, \phi_2)$ (both left or both right) consists of a smooth map $\varphi : M_1 \rightarrow M_2$ and a Lie algebra homomorphism $\rho : \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$ such that the vector fields $\phi_1(A)$ and $\phi_2 \circ \rho(A)$ are φ -related for all $A \in \mathfrak{g}_1$. If $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ and $\rho = \mathrm{id}$, φ is called equivariant, or, equivalently, a morphism of \mathfrak{g} -manifolds.

Note that the induced mapping $\mathfrak{g} \times M \rightarrow TM$, $(A, m) \mapsto (\phi(A))_m$ is automatically smooth. Using the notion of Lie algebra action, Propositions 13.4.2 and 13.4.4 may be restated as follows.

Corollary 13.4.7. 1. The Killing vector fields of a left or right Lie group action (M, G, Φ) induce a left or right action, respectively, of the Lie algebra \mathfrak{g} of G on M .
2. if (φ, ρ) is a morphism of Lie group actions, then (φ, ρ_{*e}) is a morphism of the corresponding Lie algebra actions. If φ is a morphism of G -manifolds, then it is also a morphism of the corresponding \mathfrak{g} -manifolds.

Next, consider the distribution $D^{\mathfrak{g}}$ on M spanned by the Killing vector fields, that is,

$$D_m^{\mathfrak{g}} = \{(A_*)_m \in T_m M : A \in \mathfrak{g}\}. \quad (13.4.10)$$

According to points 1 and 3 of Proposition 13.4.2, we have

$$D_{\Phi_g(m)}^{\mathfrak{g}} = (\Phi_g)_{*m} D_m^{\mathfrak{g}} \quad (13.4.11)$$

for all $g \in G$, and $\dim D_m^{\mathfrak{g}} = \dim G - \dim G_m$, respectively. In particular, $D^{\mathfrak{g}}$ need not be regular; this is so, however, if Φ is free or transitive. We will show that $D^{\mathfrak{g}}$ is integrable and that the integral manifolds are given by the orbits. For that purpose, recall the mapping $\hat{\Phi}^m : G/G_m \rightarrow M$, defined by (13.3.1).

Theorem 13.4.8 (Orbit Theorem). *Let (M, G, Φ) be a Lie group action. For every $m \in M$, $(G/G_m, \hat{\Phi}^m)$ is an initial submanifold of M whose connected components are maximal integral manifolds of $D^{\mathfrak{g}}$. In particular, $D^{\mathfrak{g}}$ is integrable.*

Proof. Assume a left action. Let $m \in M$ and let $\pi^G : G \rightarrow G/G_m$ be the natural projection. First we show that $\hat{\Phi}^m$ is an injective immersion. Injectivity was already noted before. Since $\hat{\Phi}^m \circ \pi^G = \Phi^m$, for $(\hat{\Phi}^m)_{*gG_m}$ to be injective for all $gG_m \in G/G_m$ it suffices to show that

$$\ker \Phi_{*g}^m \subset \ker \pi_{*g}^G \quad (13.4.12)$$

for all $g \in G$. To see this, let $A \in \mathfrak{g}$ be such that $\Phi_*^m A_g = 0$. Due to

$$\Phi_*^m A_g = \Phi_*^m \circ (L_g)_* A_e = (\Phi_g)_* \circ \Phi_*^m A_e = (\Phi_g)_* (A_*)_m, \quad (13.4.13)$$

and since $(\Phi_g)_*$ is bijective, we obtain $(A_*)_m = 0$ and hence $A \in \mathfrak{g}_m$ by Proposition 13.4.2(3). Then $e^{tA} \in G_m$ for all $t \in \mathbb{R}$ and hence

$$\pi_{*g}^G(A_g) = \frac{d}{dt} \Big|_{t=0} \pi^G(g e^{tA}) = 0. \quad (13.4.14)$$

This proves that $\hat{\Phi}^m$ is an injective immersion.

Secondly, we show that the connected components of $(G/G_m, \hat{\Phi}^m)$ are maximal integral manifolds for $D^{\mathfrak{g}}$. For simplicity, we may assume that G and hence G/G_m are connected. Due to π^G being a submersion, for $(G/G_m, \hat{\Phi})$ to be an integral manifold of $D^{\mathfrak{g}}$ it suffices that $\text{im}(\Phi_{*g}^m) = D_{\Phi_g(m)}^{\mathfrak{g}}$ for all $g \in G$. The latter is true because using (13.4.13) and Proposition 13.4.2(1), for $A \in \mathfrak{g}$ one finds

$$(\Phi^m)_* A_g = (\Phi^m)_* (A_*)_m = (\Phi_g)_* A_g = (\{\text{Ad}_g A\}_*)_{\Phi_g(m)} \quad (13.4.15)$$

and hence $(A_*)_{\Phi_g(m)} = (\Phi^m)_* \left(\text{Ad}_g^{-1} A \right)_m$. Maximality follows from the fact that the images of the connected components of the submanifolds $(G/G_m, \hat{\Phi}^m)$, $m \in M$, establish a disjoint decomposition of M . Finally, since m is arbitrary, we can conclude that D^g is integrable. Hence Proposition 9.5.7 yields that $(G/G_m, \hat{\Phi}^m)$ is weakly embedded. ■

As a consequence of the Orbit Theorem 13.4.8, since smooth structures of weakly embedded submanifolds are unique, the smooth structures induced on an orbit O by the submanifolds $(G/G_m, \hat{\Phi}^m)$, $m \in O$, coincide for all m , so that we can view O itself as a submanifold.

Corollary 13.4.9 ([16, Cor. 6.2.9]). *The orbits of Φ are weakly embedded (initial) submanifolds of M whose connected components are maximal integral manifolds of D^g .*

Remark 13.4.10 ([16, Rem. 6.2.10]). 1. The subspace $T_m O$ of $T_m M$ is invariant under the isotropy representation and hence the isotropy representation descends to a representation on $N_m O = T_m M / T_m O$, which is called the slice representation at m . By construction, the natural projection $T_m M \rightarrow N_m O$ intertwines the isotropy representation with the slice representation.

On the other hand, the isotropy representation restricted to $T_m O$ can be brought to the following normal form. By the Orbit Theorem 13.4.8, the map $\Phi_{*e}^m : \mathfrak{g} \rightarrow T_m M$ has image $T_m O$. By point 3 of Proposition 13.4.2 it has kernel \mathfrak{g}_m . Hence, this mapping induces a natural vector space isomorphism

$$T_m O \cong \mathfrak{g}/\mathfrak{g}_m. \quad (13.4.16)$$

In view of (13.4.2), this implies that this isomorphism intertwines the representation of G_m on $\mathfrak{g}/\mathfrak{g}_m$ induced by the adjoint representation with the isotropy representation on $T_m O$.

2. Since O is invariant under Φ and since it is a weakly embedded submanifold of M , Φ restricts to a transitive action of G on O . Since the connected components of O are integral manifolds of D^g , (13.4.11) implies that the action of G on TM induced by Φ leaves invariant the submanifold TO . The restriction of this action to the invariant submanifold $(TM)|_O$ induces an action Φ^N of G on the normal bundle $NO = (TM)|_O / TO$ by the vector bundle automorphisms (smoothness follows from the existence of local sections of the submersion $(TM)|_O \rightarrow NO$)

$$\Phi_g^N ([X]) = [(\Phi_g)_* X], \quad g \in G, X \in TM. \quad (13.4.17)$$

The natural projections $(TM)|_O / TO \rightarrow NO \rightarrow O$ are equivariant. Equivariance of the second projection implies that Φ^N restricts to an action of G_m on $N_m O$ for every $m \in O$. This action coincides with the slice representation at m .

3. A manifold carrying a transitive Lie group action is called homogeneous. The Orbit Theorem 13.4.8 yields that every homogeneous manifold is diffeomorphic to G/H for some Lie group G and some closed subgroup H of G . The diffeomorphism provided by this theorem is obviously equivariant with respect to the action of G on G/H by the induced translations \hat{L}_g (in the case of a left action) or \hat{R}_g (in the case of a right action). Hence, it is an isomorphism of Lie group actions. In this sense, all the homogeneous spaces discussed in Section 13.1 constitute homogeneous manifolds, which explains their name. The notion of homogeneous manifold may be further specialized, e.g. to that of homogeneous Riemannian manifold (a Riemannian manifold carrying a transitive Lie group action by isometries).

13.5 Proper actions

Proper group actions, which we discuss here, behave especially well. For example, they allow for local models in terms of vector bundles, which carries important implications for the structure of both the manifold carrying the action and the orbit space. Recall that a proper map, in the context of smooth manifolds, is a smooth map under which the preimages of compact subsets are compact.

Definition 13.5.1 (Proper action). A Lie group action (M, G, Φ) is called proper if the mapping

$$\Phi_{\text{ext}} : G \times M \rightarrow M \times M, \quad \Phi_{\text{ext}}(g, m) = (\Phi_g(m), m). \quad (13.5.1)$$

We give two equivalent characterizations of proper mappings. Note that since manifolds are metrizable, one may use the criterion that a subset is compact iff every sequence contains a subsequence converging in that subset.

Lemma 13.5.2 ([16, Lem. 6.3.2]). *For a continuous mapping $f : M \rightarrow N$ between smooth manifolds, TFAE:*

1. f is proper.
2. Any sequence in M whose image converges in N contains a convergent subsequence.
3. f is closed and $f^{-1}(p)$ is compact for all $p \in N$.

Proof. 1 \Rightarrow 2: Let (m_n) be a sequence in M as assumed. Since N is locally compact, there exists a compact $K \subset N$ such that $f(m_n) \in K$ for sufficiently large n . Then $m_n \in f^{-1}(K)$ for large n . Since f is proper, $f^{-1}(K)$ is compact, hence (m_n) contains a convergent subsequence.

2 \Rightarrow 3 The map f is closed: let $K \subset M$ be closed. Let (p_n) be a sequence on $f(K)$ that converges to some $p \in N$. For every n there is $m_n \in K$ such that $p_n = f(m_n)$. By assumption, (m_n) contains a subsequence converging to some $m \in M$. Since K is closed, $m \in K$. Then $p = f(m) \in f(K)$. Thus, $f(K)$ is closed. Compactness of $f^{-1}(p)$ for $p \in N$ is obvious.

$3 \Rightarrow 1$: Let $K \subset N$ be compact and let $\{U_i\}_{i \in I}$ be an open covering of $f^{-1}(K)$. Let $p \in K$. Since $\{U_i\}$ is an open covering of $f^{-1}(p)$, which is compact, there is a finite subcovering $\{U_i\}_{i \in I_p}$. Let V_p denote the complement in N of the image under f of the complement of $\bigcup_{i \in I_p} U_i$ in M . By construction, $p \in V_p$. Since f is closed, V_p is open. Hence $\{V_p\}_{p \in K}$ is an open covering of K . Since K is compact, there is a finite subcovering labelled by p_1, \dots, p_r . Then $\{U_i : i \in I_{p_1} \cup \dots \cup I_{p_r}\}$ is a finite subcovering of $\{U_i\}_{i \in I}$. ■

This immediately implies the following characterizations of proper actions.

Corollary 13.5.3 ([16, Cor. 6.3.3]). *For a Lie group action (M, G, Φ) , TFAE:*

1. *The action is proper.*
2. *If $(g_n) \in G$ and $(m_n) \in M$ are sequences such that (m_n) and $(\Phi_{g_n}(m_n))$ converge, then (g_n) contains a convergent subsequence.*
3. *The mapping Φ_{ext} is closed and all stabilizers of Φ are compact.*

Proper actions have the following elementary topological properties.

Proposition 13.5.4 ([16, Prop. 6.3.4]). *If (M, G, Φ) is a proper Lie group action, then:*

1. *The restriction of Φ to the product of a closed subgroup $H \trianglelefteq G$ with an H -invariant weakly embedded (initial) submanifold of M is proper.*
2. *The orbit map Φ^m is proper for all $m \in M$.*
3. *The orbit space M/G is Hausdorff.*

Proof. 1. We apply Corollary 13.5.3(2). $N \subset M$ and H -invariant weakly embedded submanifold. Let (g_n) and (m_n) be sequences in H and N such that (m_n) and $(\Phi_{g_n}(m_n))$ converge in N . Since N is a submanifold, the natural inclusion $N \hookrightarrow M$ is continuous. Therefore (m_n) and $(\Phi_{g_n}(m_n))$ converge in M . Since Φ is proper, (g_n) also contains a subsequence which converges in G . Since H is closed, the limit belongs to H .
2. We apply Lemma 13.5.2(2). Let $(g_n) \in G$ be a sequence such that $(\Phi^m(g_n))$ converges. Then the sequences (a_n) and $(m_n = m)$ (constant sequence) satisfy the assumptions of Corollary 13.5.3(2). Since Φ is proper, it follows that (g_n) contains a convergent subsequence.
3. The argument is analogous to that for homogeneous spaces (Lemma 13.1.1(3)). Denote the natural projection by $\pi : M \rightarrow M/G$. Let $m_1, m_2 \in M$ be such that $\pi(m_1) \neq \pi(m_2)$. Then $(m_1, m_2) \notin \text{im}(\Phi_{\text{ext}})$. Since Φ_{ext} is closed in $M \times M$, there exist open neighborhoods U_i of m_i in M such that $(U_1 \times U_2) \cap \text{im}(\Phi_{\text{ext}}) = \emptyset$. Then $\pi(U_1) \cap \pi(U_2) = \emptyset$. Since, by Proposition 13.3.3(2), the map π is open, $\pi(U_i)$ is a neighborhood of $\pi(m_i)$, $i = 1, 2$. ■

The Orbit Theorem 13.4.8 and Proposition 13.5.4(2) imply:

Corollary 13.5.5. *The orbits of a proper Lie group action are closed embedded submanifolds.*

Proof. Let $m \in M$. By Proposition 13.5.4(2) and Lemma 13.5.2(3), Φ^m is closed. Hence the image $\Phi^m(G)$ is closed and the induced map $\hat{\Phi}^m : G/G_m \rightarrow M$ is closed. Since the latter is bijective onto the orbit $\Phi^m(G)$, it is open onto $\Phi^m(G)$ in the subspace topology induced from M . ■

Remark 13.5.6. Let (M, G, Φ) be a proper Lie group action, let O be an orbit and let $m \in O$. The fact that G_m is compact has the following immediate consequences for the isotropy representation.

1. By Proposition 11.2.8, $T_m M$ admits a G_m -invariant scalar product.
2. The orthogonal complement $(T_m O)^\perp$ of $T_m O$ in $T_m M$ with respect to this scalar product is invariant under the isotropy representation. Hence the isotropy representation decomposes into the direct sum

$$T_m M \cong T_m O \oplus (T_m O)^\perp. \quad (13.5.2)$$

3. The natural projection $T_m M \rightarrow N_m O$ restricts to an isomorphism $(T_m O)^\perp \rightarrow N_m O$ which intertwines the isotropy representation induced on $(T_m O)^\perp$ with the slice representation.
4. Points 2 and 3 imply that there exists a vector space isomorphism $T_m M \cong T_m O \oplus N_m O$ intertwining that isotropy representation with the direct sum of the isotropy representation on $T_m O$ and the slice representation.

Remark 13.5.7. Let (M, Φ) and (M', Φ') be G -manifolds and let $\varphi : M \rightarrow M'$ be equivariant. If Φ' is proper, then Φ is proper, too: let (g_n) and (m_n) be sequences in G and M , respectively, such that (m_n) and $(\Phi_{g_n}(m_n))$ converge. Then $(\varphi(m_n))$ and $\Phi'_{g_n}(\varphi(m_n))$ also converge and thus (g_n) contains a convergence subsequence. This has the following consequence that *the direct product of G -manifolds is proper if one of the factors is proper*.

13.6 Equivariant tubular neighborhoods

Now we will discuss the equivariant analog of the Tubular Neighborhood Theorem 8.4.7 for the orbits of a proper Lie group action.

Definition 13.6.1 (G -vector bundle). For a Lie group G , a G -vector bundle is a **VB** with an action of G by **VB**-automorphisms. That is, a **VB** $E \xrightarrow{\pi} M$ with a Lie group action $\Phi : G \times E \rightarrow E$ such that Φ_g is a **VB**-morphism for all $g \in G$.

Definition 13.6.2 (Tubular Neighborhood and Slice). Let (M, G, Φ) be a Lie group action and let $O \subset M$ be an orbit.

1. A (equivariant) tubular neighborhood of O is an equivariant diffeomorphism $\xi : U \rightarrow E$ from an open invariant neighborhood U of O in M onto an open neighborhood V of the zero section σ_0 in a G -vector bundle E over O such that $\chi|_O = \sigma_0$.
2. For $m \in O$, the G_m -invariant neighborhood $V_m = V \cap E_m$ of the origin in E_m is called the linear slice of χ at m and the G_m -invariant embedded submanifold $U_m = \chi^{-1}(V_m)$ of M is called the slice of χ at m .

One says that the G -vector bundle E provides a local normal form for Φ near O .

Remark 13.6.3. 1. Recall from Remark 13.4.10(2) that Φ induces an action Φ^N of G on the bundle manifold of the normal bundle NO , turning NO into a G -vector bundle. We will show below that one may always choose $E = NO$.

2. Let $m \in O$. Since the orbit map $G/G_m \rightarrow O$ is an equivariant diffeomorphism, one might as well assume E to be a G -vector bundle over G/G_m . In this case the requirement $\chi|_O = \sigma_0$ reduces to $\chi(m) = \sigma(0)([e])$.
3. The slices U_m , $m \in M$, of a tubular neighborhood $\chi : U \rightarrow E$ of O have the following properties.
 - (a) $G \cdot U_m = U$ and U_m is closed in U .
 - (b) If $g \in G$ satisfies $\Phi_g(U_m) \cap U_m \neq \emptyset$, then $g \in G_m$.
 - (c) U_m intersects every orbit in U transversally.

The proof consists in showing that the fibers E_m of E have these properties (easy exercise).

The rest of this section will be devoted to proving the Equivariant Tubular Neighborhood Theorem.

Theorem 13.6.4 (Equivariant Tubular Neighborhood [16, Thm. 6.4.3]). *Every orbit of a proper Lie group action admits a (equivariant) tubular neighborhood.*

For the proof we will need a few lemmas. To be definite, we will assume that Φ is a left action. We start by constructing a mapping $\varphi : U_m \rightarrow N_m O$. According to Remark 13.5.6(1), we may choose a G_m -invariant scalar product on $T_m M$. Let $(T_m O)^\perp$ denote the corresponding orthogonal complement of $T_m O$ (although any G_m -invariant complement will do).

Lemma 13.6.5 ([16, Lem. 6.4.4]). *There exists a G_m -invariant embedded submanifold U_m of M containing m such that*

$$T_m U_m = (T_m O)^\perp \tag{13.6.1}$$

and a G_m -equivariant diffeomorphism φ from U_m onto a G_m -invariant open neighborhood of the origin in $N_m O$ such that $\varphi(m) = 0$.

Proof. ■

Next, by restriction, the action Φ induces the mapping

$$\psi : G \times U_m \rightarrow M, \quad \psi(g, \tilde{m}) = \Phi_g(\tilde{m}). \quad (13.6.2)$$

Lemma 13.6.6 ([16, Lem. 6.4.5]). *U_m can be shrunk so that*

1. ψ is a submersion,
2. If $g_1, g_2 \in G$ and $m_1, m_2 \in U_m$ satisfy $\psi(g_1, m_1) = \psi(g_2, m_2)$, then $g_2 g_1^{-1} \in G_m$ and $m_2 = \Phi_{g_2^{-1} g_1}(m_1)$.

Proof. ■

Finally, by restriction, the action Φ^N induces the mapping

$$\psi^N : G \times N_m O \rightarrow NO, \quad \psi^N(g, [X]) = \Phi_g^N([X]). \quad (13.6.3)$$

Lemma 13.6.7 ([16, Lem. 6.4.6]). 1. *The map ψ^N is a surjective submersion.*

2. If $g_1, g_2 \in G$ and $[X_1], [X_2] \in N_m O$ satisfy the relation $\psi^N(g_1, [X_1]) = \psi^N(g_2, [X_2])$, then $[X_2] = \Phi_{g_2^{-1} g_1}^N([X_1])$ and $g_2^{-1} g_1 \in G_m$.

Proof. ■

Proof of Theorem 13.6.4. We will construct a tubular neighborhood by means of the normal bundle NO and the action Φ^N . Choose a point $m \in O$. The plan of the proof can be summarized in the following commutative diagram

$$\begin{array}{ccccccc} G \times M & \xhookleftarrow{\quad} & G \times U_m & \xrightarrow{\text{id}_G \times \varphi} & G \times N_m O & \xhookrightarrow{\quad} & G \times NO \\ \Phi \downarrow & & \psi \downarrow & & \downarrow \psi^N & & \downarrow \psi^N \\ M & \xhookleftarrow{\quad} & U & \dashrightarrow_{\chi} & NO & \xrightarrow{=} & NO. \end{array} \quad (13.6.4)$$

whose ingredients were constructed step by step in the lemmas above and ultimately define the desired diffeomorphism χ .

As a consequence of Lemmas 13.6.6(2) and 13.6.7(2) and the G_m -equivariance of φ , there exists a unique injective mapping $\chi : U \rightarrow NO$ such that the central square in the diagram above commutes. Since ψ is a submersion, χ is smooth. Since ψ^N is a submersion, χ has open image V and the inverse mapping is smooth, too. Thus, χ is a diffeomorphism onto V . Since ψ intertwines the action of G on $G \times U_m$ by left translation on the first factor with Φ and since an analogous statement holds for ψ^N , χ is equivariant. This completes the proof. ■

13.7 Free proper actions

In case of a free proper action, the tubular neighborhoods constructed in Theorem 13.6.4 give us an atlas on the orbit space by means of local slices.

Corollary 13.7.1 ([16, Cor. 6.5.1]). *The orbit space M/G of a free proper Lie group action (M, G, Φ) admits a unique smooth structure such that the natural projection $M \rightarrow M/G$ is a submersion.*

Proof. Uniqueness follows from the universal property of submersions. To prove existence, let $\pi : M \rightarrow M/G$ denote the natural projection. By Proposition 13.5.4(3), the orbit space M/G is Hausdorff. By Proposition 13.3.3(3), it is second countable. Let $m \in M$ and let O be the orbit of m . Since the action Φ is free, Lemma 13.6.7 implies that the mapping $\psi_m : G \times N_m O \rightarrow NO$, defined by formula (13.6.3), is a diffeomorphism (in fact, it is an isomorphism of G -vector bundles, and so NO is globally trivial). Taking a tubular neighborhood $\chi : U \rightarrow NO$ of O and composing χ with ψ_m^{-1} , we obtain a G -equivariant diffeomorphism $\tilde{\chi}$ from U onto a G -invariant neighborhood of $G \times \{0\}$, which by G -invariance must have the form $G \times V_m$ for some open neighborhood V_m of the origin in $N_m O$. By embedding V_m as the subset $\{e\} \times V_m$ into $G \times V_m$, we obtain a bijective continuous map

$$V_m \rightarrow G \times V_m \xrightarrow{\tilde{\chi}^{-1}} U \xrightarrow{\pi} \hat{U} := \pi(U), \quad (13.7.1)$$

where $\hat{U} := \pi(U)$ is an open neighborhood of $\pi(m)$ in M/G . This map is open and hence a homeomorphism, because it maps an open subset $W \subset V_m$ to the open subset $\pi(\tilde{\chi}^{-1}(G \times W)) \subset \hat{U}$. Thus the inverse of this map defines a local chart $(\hat{U}, \hat{\kappa})$ on M/G at $\pi(m)$. The transition map between two such local charts is given by

$$\hat{\kappa}_2 \circ \hat{\kappa}_1^{-1} : \hat{\kappa}_1(\hat{U}_1 \cap \hat{U}_2) \rightarrow \hat{\kappa}_2(\hat{U}_1 \cap \hat{U}_2), \quad (13.7.2)$$

$$\hat{\kappa}_2 \circ \hat{\kappa}_1^{-1}([X]) = \text{pr}_2 \circ \tilde{\chi}_2 \circ \tilde{\chi}_1^{-1}(e, [X]). \quad (13.7.3)$$

Hence it is smooth. Thus we have constructed a smooth structure on M/G , with respect to which π is a submersion since via χ and $\hat{\kappa}$ it is locally represented by pr_2 . ■

Remark 13.7.2. This result actually holds not only for free proper actions, but for all proper actions with a single orbit type (i.e. such that the stabilizer subgroups of all points are conjugate).

Accordingly, if Φ is free and proper, M/G will be referred to as the orbit manifold of Φ . From here, we immediately get:

Corollary 13.7.3 ([16, Cor. 6.5.3]). *If (M, G, Φ) is a free proper Lie group action, then:*

1. *if $H \triangleleft G$ is a closed normal subgroup, then Φ descends to a free proper action of G/H on M/H and id_M induces a diffeomorphism between $(M/H)/(G/H)$ and M/G .*
2. *If $N \xrightarrow{\varphi} M$ is a G -invariant embedded submanifold, then the induced action of G on N is free and proper and the map $\hat{\varphi} : N/H \rightarrow M/G$, induced by passing to the quotients, defines an embedded submanifold of M/G .*

Example 13.7.4 ([16, Example 6.5.4]). 1. The action of $G = \mathbb{Z}$ on $M = \mathbb{R} \times (-1, 1)$ defined by $\Phi_k(x, t) = (x + k, (-1)^k t)$ is free. It is also proper (recall that a proper action of a discrete group is called properly discontinuous): it is easy to see that if (x_n, t_n) and $\Phi_{k_n}(x_n, t_n)$ both converge as $n \rightarrow \infty$, then k_n must become constant at large n , hence it converges as well. The quotient manifold coincides with the Möbius strip (the smooth structure coincides because the natural projection is a smooth submersion with respect to either definition of the Möbius strip and submersions have a universal property). Applying Corollary 13.7.3 to the closed normal subgroup $H = 2\mathbb{Z}$ yields that the Möbius strip is diffeomorphic to the quotient manifold of the free proper action of $G/H = \mathbb{Z}_2$ on $M/H \cong \mathbb{S}^1 \times (-1, 1)$, where the generator of \mathbb{Z}_2 acts by $(\alpha, t) \mapsto (-\alpha, -t)$, as expected.

2. Let \mathbb{K}_1 be the Lie group given by the unit sphere in $\mathbb{K} = \mathbb{R}, \mathbb{C}$, or \mathbb{H} (so $\mathbb{R}_1 = O_1$, $\mathbb{C}_1 = U_1$, and $\mathbb{H}_1 = \text{Sp}_1 \cong \text{SU}_2$). The action of \mathbb{K}_1 on \mathbb{K}^n is proper because \mathbb{K}_1 is compact. Since this action is isometric with respect to the standard dot product, it restricts to an action of \mathbb{K}_1 on the unit sphere \mathbb{S}^{dn-1} , where $d = \dim_{\mathbb{R}} \mathbb{K}$. The latter action is in addition free. The quotient manifold $\mathbb{S}^{dn-1}/\mathbb{K}_1$ coincides with the projective space $\mathbb{K}P^{n-1}$. In the case of U_1 acting on \mathbb{C}^2 this construction reproduces exactly the Hopf bundle $\mathbb{S}^1 \hookrightarrow \mathbb{S}^3 \rightarrow \mathbb{S}^2$.

Corollary 13.7.1 gives us a very important class of fiber bundles, which we now turn to. The Tubular Neighborhood Theorem guarantees that every free proper action generates equivariant local trivializations of $\pi : M \rightarrow M/G$ by neighborhoods of the form $U_m \times G$, which makes it a fiber bundle. Conversely, a free action that admits such local trivializations must be proper:

Proposition 13.7.5. *Let (P, G, Φ) be a free action Lie group action and let $\pi : P \rightarrow M$ be a smooth map such that for every $m \in M$ there is an equivariant local trivialization at m , that is, an open neighborhood W of m and a diffeomorphism $\chi : \pi^{-1}(W) \rightarrow W \times G$ such that*

1. *χ intertwines Φ with the G -action on $W \times G$ by translations (left or right according to the type of Φ) on the factor G ,*
2. *$\text{pr}_1 \circ \chi(p) = \pi(p)$ for all $p \in \pi^{-1}(W)$.*

Then Φ is proper, M is diffeomorphic to the orbit manifold P/G , and π corresponds, via this diffeomorphism to the natural projection $P \rightarrow P/G$.

Proof. Exercise. ■

Denoting $\kappa = \text{pr}_2 \circ \chi : \pi^{-1}(W) \rightarrow G$, for a right action Φ , property 1 can be rewritten as

$$\kappa(\Phi_g(p)) = \kappa(p)g, \quad p \in \pi^{-1}(W), \quad a \in G. \quad (13.7.4)$$

Definition 13.7.6 (Principal fiber bundle I). A **principal fiber bundle (PFB)** is the fiber bundle $\pi : P \rightarrow P/G$ associated with a free proper Lie group action (P, G, Φ) . Equivalently, it is a tuple $(P \xrightarrow{\pi} M, G, \Phi)$, where (P, G, Ψ) is a free Lie group action and π is a smooth map that admits equivariant local trivializations at every $m \in M$ in the sense of Proposition 13.7.5. The group G is called the structure group of P , and P is also called a principal G -bundle. The existence of local trivializations implies that π is a submersion. Hence the fibers of P are submanifolds, each of them is diffeomorphic to G , and in fact is a principal homogeneous G -manifold.

Example 13.7.7 (Homogeneous spaces as principal bundles). Every homogeneous space G/H , where $H \trianglelefteq G$ is a closed subgroup, is a principal H -bundle because the action of H on G is always free and proper. E.g. if $\Phi : H \times G \rightarrow G$ is the action by right translations, then $\Phi_{\text{ext}}(h, g) = (g, gh)$, which is clearly a closed map, hence proper by Proposition 13.5.3(3).

Example 13.7.8. Every Stiefel manifold $V_k(\mathbb{K}^n)$ (cf. 13.1.8) is therefore the base manifold of a O_k -bundle with total space O_n . A particular instance of an orthogonal Stiefel bundle is

$$\pi : \text{SO}_3 \rightarrow V_1(\mathbb{R}^3) \cong \mathbb{S}^2. \quad (13.7.5)$$

We can see that this bundle is in fact isomorphic to the unit circle bundle of $T\mathbb{S}^2$:

$$S(T\mathbb{S}^2) = \{(\mathbf{x}, \mathbf{v}) \in \mathbb{S}^2 \times \mathbb{S}^3 \mid \mathbf{x} \perp \mathbf{v}\}. \quad (13.7.6)$$

Indeed, to each point (\mathbf{x}, \mathbf{v}) in this bundle we can assign a unique positively oriented orthonormal frame $(\mathbf{x}, \mathbf{v}, \mathbf{x} \times \mathbf{v})$, i.e. an element of SO_3 . This is an isomorphism of SO_3 -torsors $\text{SO}_3 \cong S(T\mathbb{S}^2)$, and in particular a vertical isomorphism of fiber bundles over \mathbb{S}^2 . Therefore the principal U_1 -bundle $\text{SO}_3 \rightarrow \mathbb{S}^2$ is isomorphic to the circle bundle of the complex line bundle $O(2)$ (see Example 8.2.2), and has Chern number 2.

Another Stiefel bundle is the principal U_1 -bundle

$$\pi' : \text{SU}_2 \rightarrow V_1(\mathbb{C}^2) \cong \mathbb{C}P^1. \quad (13.7.7)$$

13.8 Example: flag manifolds

Definition 13.8.1 (Grassmann manifolds). The Grassmann manifold $G_k(\mathbb{K}^n)$ is defined as the set of k -dimensional subspaces of \mathbb{K}^n . As for the Stiefel manifolds, the manifold structure is obtained by identifying it with a homogeneous space. Any k -dimensional subspace of \mathbb{K}^n can be obtained from the subspace spanned by the first k elements of the standard basis by application of a linear isometry. This transformation is determined by the given subspace up to prior application of a block matrix

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad (13.8.1)$$

where a and b are isometries of \mathbb{K}^k and \mathbb{K}^{n-k} , respectively. Hence,

$$G_k(\mathbb{R}^n) = O_n / (O_k \times O_{n-k}), \quad (13.8.2)$$

$$G_k(\mathbb{C}^n) = U_n / (U_k \times U_{n-k}), \quad (13.8.3)$$

$$G_k(\mathbb{H}^n) = Sp_n / (Sp_k \times Sp_{n-k}), \quad (13.8.4)$$

where the subgroups consist of block-diagonal matrices of the form (13.8.1). The case $k = 1$ reproduces the projective spaces $G_1(\mathbb{K}^n) = \mathbb{K}P^{n-1}$ (this is just a homeomorphism for now; to show that this is a diffeomorphism we will need to study more general Lie group actions on manifolds and identify $G_k(\mathbb{K}^n)$ with the orbit manifold of the action of the group of isometries of \mathbb{K}^{n-k} on $V_k(\mathbb{K}^n)$).

Definition 13.8.2 (Infinite Grassmannians). Similarly to the case of infinite (or *stable*) Stiefel spaces above, we can define the direct limits

$$G_k(\mathbb{K}^\infty) = \bigcup_n G_k(\mathbb{K}^n). \quad (13.8.5)$$

Let us denote by $O_n = O_n(\mathbb{K})$ for $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ the corresponding group of isometries (O_n , U_n , and Sp_n , respectively). Consider the smooth left action of O_n on $\mathbb{K}^{n \times k}$ given by left multiplication $(g, A) \mapsto gA$. The orbit of the element $p = (e_1, \dots, e_k)$, where $e_1, \dots, e_n \in \mathbb{K}^n$ is the standard basis, is diffeomorphic to the Stiefel manifold $V_k(\mathbb{K}^n)$. As we have seen, the quotient map $\alpha_n : O_n \rightarrow V_k(\mathbb{K}^n)$ given by $\alpha_n(g) = (ge_1, \dots, ge_k)$ is a locally trivial fiber bundle. The stabilizer of the element p is nothing by the subgroup $O_{n-k} \subset O_n$ which acts as identity on the first k components of \mathbb{K}^n . We shall denote this subgroup by O_{n-k} . The inclusions $\mathbb{K}^n \hookrightarrow \mathbb{K}^{n+1}$ induce the compatible inclusions of Stiefel manifolds $V_k(\mathbb{K}^n) \hookrightarrow V_k(\mathbb{K}^{n+1})$ and the corresponding principal bundles

$$O_k \hookrightarrow O_n \rightarrow V_k(\mathbb{K}^n), \quad (13.8.6)$$

as well as the (topological) principal bundle

$$O_k \hookrightarrow O_\infty \rightarrow V_k(\mathbb{K}^\infty). \quad (13.8.7)$$

over the infinite Stiefel space of k -frames in \mathbb{K}^∞ .

Now, each element of $V_k(\mathbb{K}^n)$ can be identified with a k -frame, and thus determines the linear k -dimensional subspace $W \subset \mathbb{K}^n$ spanned by that frame. Thus we have a quotient map $V_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n)$. The fiber of this map above W consists of all orthonormal frames on W and thus is identified with the orbit of the free and proper action of O_k on $V_k(\mathbb{K}^n)$. Thus $V_k(\mathbb{K}^n)$ becomes a principal bundle over the Grassmann manifold (we shall call it the *Stiefel principal bundle*):

$$O_k \hookrightarrow V_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n), \quad (13.8.8)$$

and by taking the direct limit with respect to the inclusions $\mathbb{K}^n \hookrightarrow \mathbb{K}^{n+1}$ we also get the (topological) principal bundle

$$O_k \hookrightarrow V_k(\mathbb{K}^\infty) \rightarrow G_k(\mathbb{K}^\infty). \quad (13.8.9)$$

We thus have the alternative definition of the Grassmann manifolds not as homogeneous spaces, but as quotients of the Stiefel manifolds by a proper free action

$$G_k(\mathbb{K}^n) \cong V_k(\mathbb{K}^n)/O_k, \quad G_k(\mathbb{K}^\infty) \cong V_k(\mathbb{K}^\infty)/O_k \quad (13.8.10)$$

(in the infinite case we didn't prove any theorems about quotients by a free group action, but this issue can be circumvented by identifying the quotient with the direct limit of the finite-dimensional smooth ones).

Remark 13.8.3. Define the square matrix $\tau = \begin{bmatrix} 0 & I_k \\ I_{n-k} & 0 \end{bmatrix}$. Then the conjugation $A \mapsto \tau A \tau^T$ on O_n takes the subgroup $O_k \times O_{n-k}$ to $O_{n-k} \times O_k$ and hence induces a diffeomorphism $G_k(\mathbb{K}^n) \cong G_{n-k}(\mathbb{K}^n)$. Therefore, for the purposes of studying the topology of these spaces, we can always assume $n \geq 2k$.

Proposition 13.8.4 ([39, Thm. 10.9.5]). *If $n \geq 2k$, we have*

$$\pi_q(G_k(\mathbb{K}^n)) \cong \pi_q(V_k(\mathbb{K}^n)) \times \pi_{q-1}(O_k), \quad q \geq 1. \quad (13.8.11)$$

Proof. If $n \geq 2l$, then the homotopy

$$(A, t) \mapsto \begin{bmatrix} A \cos \frac{\pi t}{2} \\ I_k \sin \frac{\pi t}{2} \\ 0 \end{bmatrix}, \quad 0 \leq t \leq 1, \quad (13.8.12)$$

shows that the inclusion map $O_k \hookrightarrow V_k(\mathbb{K}^n)$ is nullhomotopic. Then the long exact sequence of homotopy groups (Theorem 5.2.9) for the fibration $O_k \hookrightarrow V_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n)$ implies the asserted isomorphism. ■

Thus we can compute the homotopy groups of Grassmann manifolds provided we compute those of Stiefel manifolds and orthogonal groups.

Proposition 13.8.5 (Stable homotopy groups [39, Thm. 10.9.6]). *The homomorphism $\pi_q(O_n) \rightarrow \pi_q(O_{n+1})$ induced by the inclusion $O_n \hookrightarrow O_{n+1}$ is an isomorphism for $q < d(n+1) - 2$ and an epimorphism for $q = d(n+1) - 2$. Equivalently, the pair (O_{n+1}, O_n) is $(d(n+1) - 2)$ -connected.*

Proof. Let us consider the standard action of O_{n+1} on the unit sphere $\mathbb{S}^{d(n+1)-1} \subset \mathbb{K}^{n+1}$, where $d = \dim_{\mathbb{R}} \mathbb{K}$. Given any $v \in \mathbb{S}^{d(n+1)-1}$, there exists $A \in O_{n+1}$ whose last column is v , i.e. $Ae_{n+1} = v$. This means that the action is transitive, and clearly the stabilizer of e_{n+1} is the subgroup O_n . This we have the PFB

$$O_n \hookrightarrow O_{n+1} \rightarrow \mathbb{S}^{d(n+1)-1}. \quad (13.8.13)$$

The long exact sequence of this fibration combined with the triviality of the homotopy groups of spheres up to $d(n+1) - 2$ gives the asserted result. ■

Corollary 13.8.6. *The pairs (SO_{n+1}, SO_n) are $(n-1)$ -connected (by using the fibration $SO_n \xrightarrow{\det} O_n \xrightarrow{\det} \{\pm 1\}$).*

Remark 13.8.7. Thus fixing an integer $q > 0$, we choose $n \geq \lceil (q+2)/d \rceil$. It follows that the homomorphism $\pi_q(O_n) \rightarrow \pi_q(O_\infty)$ is an isomorphism for all such n . We say that the homotopy group $\pi_q(O_n)$ stabilizes after $n = \lceil (q+2)/d \rceil$. This is called the stable range. For example, the stable range for π_q of real orthogonal groups is $[q+2, \infty)$.

In the case of unitary groups we have $d = 2$, so the isomorphism proven above holds for $q < 2n$. In particular, since we know all of the homotopy groups of $U_1 \cong \mathbb{S}^1$ and $SU_1 = \{1\}$, this is enough to compute π_1 and π_2 of all unitary groups. Moreover, we know that $\pi_3(SU_2) = \pi_3(\mathbb{S}^3) \cong \mathbb{Z}$, and $3 < 2 \cdot 2$, so this lets us compute π_3 of unitary groups as well.

Corollary 13.8.8 (π_q of unitary groups). $\pi_1(SU_n) = 0$, $\pi_1(U_n) \cong \mathbb{Z}$, $\pi_2(SU_n) = \pi_2(U_n) = 0$, $\pi_3(SU_n) \cong \pi_3(U_n) \cong \mathbb{Z}$.

Proof. Take the last Proposition and set $i = 1$. This gives an isomorphism $\pi_1(O_n) \cong \pi_1(O_{n+1})$. Since $\pi_1(SU_1)\pi_1(\{1\}) = 0$ and $\pi_1(U_1) = \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$, this proves the first two equalities.

The same principle applied to $i = 2$ and $\pi_2(\mathrm{SU}_1) = \pi_2(\mathrm{U}_1) = 0$ gives $\pi_2(\mathrm{SU}_n) = \pi_2(\mathrm{U}_n) = 0$.

Finally, for $i = 3$ we use the fibration $\mathrm{SU}_n \hookrightarrow \mathrm{U}_n \xrightarrow{\det} \mathbb{S}^1$ to first conclude that $\pi_1(\mathrm{SU}_n) \cong \pi_q(\mathrm{U}_n)$ for all $q \geq 2$. Thus $\pi_3(\mathrm{U}_2) \cong \pi_3(\mathrm{SU}_2) \cong \mathbb{Z}$ and by the same argument as above get the same for all higher unitary groups. ■

- Remark 13.8.9.**
1. Note that in fact the determinant fibration used in this proof is trivial: $\mathrm{U}_n \cong \mathrm{SU}_n \times \mathbb{S}^1$. Indeed, the trivialization is given by $A \mapsto (\mathrm{diag}(\det A^{-1}, 1, \dots, 1) \cdot A, \det A)$. Thus instead of the exact sequence we could also use $\pi_q(\mathrm{U}_n) \cong \pi_q(\mathrm{SU}_n) \times \pi_q(\mathbb{S}^1) = \pi_q(\mathrm{SU}_n)$ for $q > 1$.
 2. We will later prove, using Morse theory, the very deep fact that $\pi_2(G) = 0$ for *any* connected Lie group G .

Corollary 13.8.10. *The pair (O_∞, O_n) is $(d(n+1) - 2)$ -connected.*

Corollary 13.8.11. $\pi_q(V_r(\mathbb{K}^{n+r})) = 0$ for $q \leq d(n+1) - 2$.

Proof. This follows from the exact sequence of the fibration $O_n \hookrightarrow O_{n+r} \rightarrow V_r(\mathbb{K}^{n+r})$ combined with the preceding Proposition. ■

Corollary 13.8.12. *By taking the limit $n \rightarrow \infty$, we conclude $\pi_i(V_k(\mathbb{K}^\infty)) = 0$ for all i .*

Remark 13.8.13. It is in fact possible to provide a CW-structure for Stiefel manifolds, including the infinite one, so by Whitehead's Theorem we can conclude that $V_k(\mathbb{K}^\infty)$ is contractible. Below we provide a direct proof. The fact that this gives us a contractible principal O_k -bundle will be crucial later when we study the classification of principal bundles.

Let us provide an explicit deformation retraction here, extending the one we constructed for \mathbb{S}^∞ (which corresponded to the case $k = 1$).

Proposition 13.8.14. *$V_k(\mathbb{K}^\infty)$ is contractible.*

Proof. Take the homotopy $h_t : \mathbb{K}^\infty \rightarrow \mathbb{K}^\infty$ defined in (5.3.10). It is linear for each t and its kernel is trivial. Thus, if we apply h_t to a k -frame, we get a k -tuple of linearly independent vectors, which can be made orthogonal by the Gram-Schmidt process. Thus we have a deformation retraction of $V_k(\mathbb{K}^\infty)$ onto the subspace of k -frames with first coordinate zero. Iterating this k times, we deform into the subspace of k -frames with first k coordinates zero. For such a k -frame (v_1, \dots, v_k) define a homotopy

$$(1-t)(v_1, \dots, v_k) + t(e_1, \dots, e_k) \tag{13.8.14}$$

where e_i is the i -th standard basis vector in \mathbb{K}^∞ . Since the first k coordinates of all v_i 's vanish, this homotopy preserves linear independence, so after again applying Gram-Schmidt we have a deformation *through* k -frames, which finishes the construction of a contraction of $V_k(\mathbb{K}^\infty)$. ■

By combining this with Proposition 13.8.4 (or directly from the long exact sequence of the infinite principal bundle), we get:

Corollary 13.8.15. $\pi_q(O_k) \cong \pi_{q+1}(G_k(\mathbb{K}^\infty))$, $q \geq 1$.

The isomorphisms between homotopy groups of the spaces $O_n \cong V_n(\mathbb{K}^n)$ (Proposition 13.8.5) also extend to more general Stiefel manifolds.

Corollary 13.8.16. Consider the inclusion $j : V_k(\mathbb{K}^n) \hookrightarrow V_{k+1}(\mathbb{K}^{n+1})$ given by $(v_1, \dots, v_k) \mapsto (v_1, \dots, v_k, e_{n+1})$. Then the induced homomorphism $j_\# : \pi_i(V_k(\mathbb{K}^n)) \rightarrow \pi_i(V_{k+1}(\mathbb{K}^{n+1}))$ is an isomorphism for $i \leq d(n+1)-2$ and an epimorphism for $i = d(n+1)-1$.

Proof. This follows from the fact that j is actually the inclusion of the fiber over e_{n+1} in the fibration

$$V_k(\mathbb{K}^n) \hookrightarrow V_{k+1}(\mathbb{K}^{n+1}) \xrightarrow{q} \mathbb{S}^{d(n+1)-1} \quad (13.8.15)$$

given by the projection q onto the last coordinate vector. The long exact sequence of homotopy groups combined with the triviality of the homotopy groups of spheres in the specified range gives the result. ■

For real orthogonal groups $d = 1$ and thus the isomorphism $\pi_q(O_n) = \pi_q(O_{n-1})$ holds only for $q < n - 2$, to the smallest group we can reduce this to is O_{q+3} . Luckily we already know that there is a homeomorphism $SO_4 \cong SO_3 \times \mathbb{S}^3$ (Example 6.4.3), and of course $SO_3 \cong \mathbb{RP}^3$. This lets us compute at least the fundamental groups.

Corollary 13.8.17. $\pi_1(O_n) \cong \pi_1(SO_n) \cong \mathbb{Z}_2$ for $n \geq 3$. Also $\pi_2(SO_n) = \pi_2(O_n) = 0$ for $n \leq 4$, and $\pi_3(SO_4) \cong \pi_3(O_4) \cong \mathbb{Z}^2$.

Proof. First we note that O_n consists of two connected components, each diffeomorphic to SO_n , therefore it suffices to compute the homotopy groups $\pi_q(SO_n)$ for $q > 0$. Now we use the homeomorphism $SO_4 \cong SO_3 \times \mathbb{S}^3$ (i.e. that the principal SU_2 -bundle $SU_2 \hookrightarrow SO_4 \rightarrow SO_3$ described in Example 6.4.3 is trivial) to compute

$$\pi_q(SO_4) \cong \pi_q(SO_3 \times \mathbb{S}^3) \cong \pi_q(\mathbb{RP}^3) \times \pi_q(\mathbb{S}^3). \quad (13.8.16)$$

Now we note that \mathbb{S}^3 is the universal covering space of \mathbb{RP}^3 , so we have $\pi_1(\mathbb{RP}^3) = \mathbb{Z}_2$ and these two spaces have isomorphic higher homotopy groups π_q , $q > 1$. Thus, using

$\pi_1(\mathbb{S}^3) = \pi_2(\mathbb{S}^3) = 0$ and $\pi_3(\mathbb{S}^3) \cong \mathbb{Z}$, we find the asserted results for $\pi_2(\mathrm{SO}_3)$, $\pi_2(\mathrm{SO}_4)$, and $\pi_3(\mathrm{SO}_4)$. Finally, the iterative isomorphism from above leads to

$$\pi_1(\mathrm{SO}_n) \cong \pi_1(\mathrm{SO}_4) \cong \pi_1(\mathbb{R}P^3) \cong \mathbb{Z}_2. \quad (13.8.17)$$

■

Remark 13.8.18. As with unitary groups, actually $\pi_2(\mathrm{SO}_n) = 0$ for all n , but the method we used above lets us reduce this only to the problem of proving $\pi_2(\mathrm{SO}_5) = 0$. Similarly, $\pi_3(\mathrm{SO}_n)$ stabilizes starting with $n = 6$, so we would need to compute $\pi_3(\mathrm{SO}_6) = \mathbb{Z}$ (and in fact $\pi_3(\mathrm{SO}_5) = \mathbb{Z}$ as well). We will not attempt this here, but these groups can be computed using the CW-structure on SO_n described in Section 6.5.

Each Stiefel O_k -bundle $V_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n)$ comes with a closely related (we will see how later) vector bundle.

Definition 13.8.19 (Tautological vector bundles). Let $G_k(V)$ be the Grassmann manifold of the vector space $V = \mathbb{K}^n$. The tautological vector bundle $\gamma_k(V) \rightarrow G_k(V)$ is defined as the following subbundle of the trivial bundle $G_k(V) \times V$:

$$\gamma_k(V) = \{(W, v) \in G_k(V) \times V : v \in W\}, \quad (13.8.18)$$

i.e. the fiber over $W \in G_k(V)$ is W itself.

Proposition 13.8.20. *The tautological bundle $\gamma_k(V)$ is a smooth VB.*

Proof. Since we already know that the Stiefel bundle $V_k(V) \rightarrow G_k(V)$ is a smooth bundle, it has local smooth sections. Consider a neighborhood $U \subset G_k(V)$ that admits such a local section $\sigma : U \rightarrow V_k(V)$. Its value at point $W \in G_k(V)$ is an orthonormal k -frame that spans the space W . This means that $\sigma(W)$ can be identified with an element of V^k . But then by projecting onto one of the k components of V^k we get k smooth local sections of $G_k(V) \times V$ which, at each point $W \in U$, happen to span exactly the space W . By Theorem 8.1.21(3) the existence of a smooth frame is equivalent to the existence of a smooth local trivialization. ■

Finally, we define an even further generalization of Grassmann manifolds, although it will not be as useful to us.

Definition 13.8.21 (Flags, flag manifolds). Let V be a \mathbb{K} -vector space of dimension $n > 1$ and let $0 < k_1 < \dots < k_{r-1} < n$ be integers. A flag in V of type $K = (k_1, \dots, k_{r-1})$ is a nested sequence of linear subspaces $V_1 \subset V_2 \subset \dots \subset V_{r-1} \subset V$ with $\dim W_i = k_i$ for each i . The set of all flags of type K in V is denoted by $F_K(V)$. $\mathrm{GL}(V)$ acts transitively on $F_K(V)$. By picking a dot product on V , setting $W_1 = V_1$ and defining W_{i+1} to be the

orthogonal complement of V_i inside V_{i+1} , with every flag we can associate a decomposition $V = W_1 \oplus \cdots \oplus W_r$ into mutually orthogonal subspaces, and this defines a bijection from $F(V)$ onto the set of orthogonal direct sum decompositions of V into subspaces of dimensions (n_1, \dots, n_r) , where $n_1 = k_1$, $n_r = n - k_{r-1}$, and $n_{i+1} = k_{i+1} - k_i$ in between. An argument similar to that for the Grassmann manifolds shows that

$$F_K(V) \cong O_n / (O_{n_1} \times \cdots \times O_{n_r}), \quad (13.8.19)$$

where the subgroups consist of block-diagonal matrices. As before, these equalities define a smooth structure on $F_K(V)$.

The case $r = 1$ and $n_1 = k$ reproduces the Grassmann manifolds $F_{(k,n-k)}(\mathbb{K}^n) \cong G_k(\mathbb{K}^n)$.

13.9 Example: projective, conformal, and CR spheres

The same homogeneous space can arise from different Lie groups G . Each of these Lie groups induces a corresponding action on that space, and a certain canonical G -invariant structure. The constructions of the sphere \mathbb{S}^n correspond to the invariant Riemannian metric of the “round sphere” inherited from the standard O_n -action. In this section we present a few other ways the sphere can be represented as a homogeneous space with corresponding invariant structures.

Definition 13.9.1 (Projective space and sphere). Consider the tautological bundle $\gamma_1(\mathbb{R}^{n+1}) \rightarrow \mathbb{R}P^n$. Its total space can be identified with $\mathbb{R}^{n+1} \setminus \{0\}$ and the structure group is \mathbb{R}^\times . The standard action of $SL_{n+1}(\mathbb{R})$ on \mathbb{R}^{n+1} restricts to an action on $\gamma_1(\mathbb{R}^{n+1})$. The resulting diffeomorphisms of $\mathbb{R}P^n$ are exactly the projective transformations. The stabilizer subgroup of the line through $e_1 \in \mathbb{R}^{n+1}$ is given by

$$P = \left\{ \begin{pmatrix} \det(A)^{-1} & \phi \\ 0 & A \end{pmatrix}, A \in GL_n(\mathbb{R}), \phi^T \in \mathbb{R}^n \right\}. \quad (13.9.1)$$

This is an example of a parabolic subgroup of the simple Lie group $SL_{n+1}(\mathbb{R})$. The identity component P_0 (with $\det A > 0$) is the stabilizer of the ray spanned by e_1 . The corresponding homogeneous space of rays in \mathbb{R}^{n+1} can be identified with \mathbb{S}^n . Thus

$$\mathbb{R}P^n \cong SL_{n+1}(\mathbb{R})/P, \quad \mathbb{S}^n \cong SL_{n+1}(\mathbb{R})/P_0. \quad (13.9.2)$$

The $SL_{n+1}(\mathbb{R})$ -invariant structure on $\mathbb{R}P^n$ associated with this principal bundle is the family of all projective lines in $\mathbb{R}P^n$ since projective lines are exactly the images under the projection $\mathbb{R}^{n+1} \setminus \{0\} \rightarrow \mathbb{R}P^n$ of the intersections of planes through the origin with $\mathbb{R}^{n+1} \setminus \{0\}$.

The idea of getting a transitive action on $\mathbb{R}P^n$ (resp. \mathbb{S}^n) by viewing it as the space of lines

(resp. rays) in \mathbb{R}^{n+1} clearly generalizes to certain subgroups of $\mathrm{SL}_{n+1}(\mathbb{R})$. If we take the sphere and an orthogonal group, then we return to the setting of the previous section, since then the embedding of \mathbb{S}^n as the unit sphere is equivariant, so we can view it as a subspace as well. Hence, in the realm of simple groups the main remaining possibility is to consider the group $\mathrm{Sp}_{2n}(\mathbb{R})$ of transformations preserving the standard symplectic form on \mathbb{R}^{2n} .

Definition 13.9.2 (Contact spheres). The linear action of $\mathrm{Sp}_{2n}(\mathbb{R})$ induces transitive actions on S^{2n-1} and \mathbb{RP}^{2n-1} . The stabilizer subgroup P of the line through e_1 again turns out to be a parabolic subgroup of $\mathrm{Sp}_{2n}(\mathbb{R})$, and its identity component P_0 is the stabilizer of the ray through e_1 .

Our next example is certainly the best studied example of a parabolic geometry, namely conformal structures. Consider \mathbb{R}^{n+2} endowed with an inner product $\langle \cdot, \cdot \rangle$ of signature $(n+1, 1)$, and the *light cone* C of nonzero null vectors, i.e. $C = \{v \in \mathbb{R}^{n+2} : v \neq 0, \langle v, v \rangle = 0\}$. Taking the standard inner product of this signature, C consists of all nonzero solutions of the equation $x_{n+2}^2 = x_1^2 + \cdots + x_{n+1}^2$. In particular, any element of C has nonzero last coordinate. Consider the space of lines in C , i.e. C/\sim where $x \sim y$ iff $x = \lambda y$ for some $\lambda \in \mathbb{R}^\times$. Viewing \mathbb{S}^n as the unit sphere in the subspace $\mathbb{R}^{n+1} \subset \mathbb{R}^{n+2}$, the map $x \mapsto (x, 1)$ includes \mathbb{S}^n into C and thus induces a smooth map to the quotient space C/\sim . Conversely, for $v = (v_1, \dots, v_{n+2}) \in C$ we get $v_{n+2} \neq 0$ and $\frac{1}{v_{n+2}}(v_1, \dots, v_{n+1})$ lies in $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. This defines a smooth map $C \rightarrow \mathbb{S}^n$ which factors to an inverse $C/\sim \rightarrow \mathbb{S}^n$ of the map $\mathbb{S}^n \rightarrow C/\sim$ above. Hence we may view \mathbb{S}^n as the space of null lines in $(\mathbb{R}^{n+2}, \langle \cdot, \cdot \rangle)$. This immediately suggests another way to make \mathbb{S}^n into a homogeneous space.

Definition 13.9.3 (Conformal sphere). Consider the standard action of $O_{n+1,1}$ on \mathbb{R}^{n+2} which leaves the null cone C invariant. Since it acts linearly, it factors to a smooth action on $\mathbb{S}^n \cong C/\sim$. In fact the action on C is transitive and this we get a diffeomorphism

$$\mathbb{S}^n \cong O_{n+1,1}/P, \quad (13.9.3)$$

where P is the stabilizer subgroup of some point of \mathbb{S}^n . The subgroup P is usually called the Poincaré conformal group. For now we just note that P is parabolic and is the semidirect product of the conformal group CO_n with an n -dimensional vector group.

Clearly the Riemannian metric induced on this sphere \mathbb{S}^n is only determined up to overall scale, and $O_{n+1,1}$ acts on \mathbb{S}^n by diffeomorphisms that preserve the standard conformal structure on \mathbb{S}^n . One may also replace $O_{n+1,1}$ with $\mathrm{SO}_{n+1,1}$ or $\mathrm{Spin}_{n+1,1}(\mathbb{R})$. However, there is no analog to the flat Euclidean space, i.e. conformal geometry does not allow a homogeneous model isomorphic to \mathbb{R}^n . On \mathbb{R}^n there are certain local conformal transformations which do not globalize.

The final example is the Hermitian analog of the conformal sphere, which is also a parabolic

contact structure.

Example 13.9.4 (CR-sphere). Consider \mathbb{C}^{n+2} with the Hermitian inner product $\langle \cdot, \cdot \rangle$ of signature $(n+1, 1)$. Let C be the cone of nonzero null vectors, which is a submanifold of real codimension 1, and consider the quotient C/\sim by the action of \mathbb{C}^* . Similar to the last example, C/\sim is diffeomorphic to \mathbb{S}^{2n+1} . The standard action of $SU_{n+1,1}$ on \mathbb{C}^{n+2} leaves C invariant and descends to a transitive action on \mathbb{S}^{2n+1} with the stabilizer of a point isomorphic to a parabolic subgroup P . Thus

$$\mathbb{S}^{2n+1} \cong SU_{n+1,1}/P. \quad (13.9.4)$$

The $SU_{n+1,1}$ -invariant structure on the odd-dimensional sphere induced by this construction is the so called almost CR-structure, which consists of a Hermitian metric identified only up to a factor $\lambda \in \mathbb{C}^*$ and a complex structure J . It also determines a unique contact structure.

13.10 (*) Diffeomorphism groups

Much of basic differential geometry on manifolds that we've covered translates to the locally convex case: if M is modeled on a locally convex space V , then its tangent bundle TM (with tangent vectors defined as equivalence classes of curves at a point) has smooth local trivializations over $U_\alpha \subset M$ mapping to $U_\alpha \times V$, and the transition maps are smooth as functions of $(m, v) \in U_{\alpha\beta} \times V$. We can construct the functor T by defining derivatives of smooth maps between locally convex manifolds. Since the pairing between V and V^* is usually not very nice, one often finds another vector space W (which can be thought of as a subspace of V^*) with a nondegenerate pairing $V \times W \rightarrow \mathbb{K}$. This allows one to define natural adjoints of linear maps and hence a cotangent bundle T^*M with typical fiber W . Smooth vector fields can be defined as usual, and $\mathfrak{X}(M)$ is still a Lie algebra with the Lie bracket. The Lie bracket is again natural with respect to smooth maps. The calculus of differential forms then also extends once we've defined T^*M .

For integration of differential forms one requires Mackey completeness. Then e.g. the Stokes theorem for differential forms on manifolds with boundaries holds. The Poincaré Lemma holds in the following form.

Lemma 13.10.1 (Poincaré Lemma [34, Lem. II.4.11]). *Let V be locally convex, W Mackey complete locally convex, and $U \subset V$ an open star-shaped (w.r.t 0) subset. Let $\omega \in \Omega^{k+1}(U, W)$ be a W -valued closed $(k+1)$ -form. Then $\omega = d\varphi$ for some $\varphi \in \Omega^k(U, W)$ satisfying $\varphi(0) = 0$ given by*

$$\varphi(x)(v_1, \dots, v_k) = \int_0^1 t^k \omega(tx)(x, v_1, \dots, v_k) dt. \quad (13.10.1)$$

To prove further results such as de Rham's Theorem, one usually needs partitions of unity, which don't always exist even on Banach manifolds. Thus a manifold M is called smoothly paracompact if every open cover has a subordinated smooth POU. De Rham's Theorem holds on such Fréchet manifolds. Examples of Banach spaces which are not smoothly paracompact are $C([0, 1], \mathbb{R})$ and $\ell^1(\mathbb{N}, \mathbb{R})$. On these spaces there exist no non-zero smooth functions supported on the unit ball.

Many familiar results about Lie groups now translate to the locally convex case. The Lie algebra of a locally convex Lie group is obviously locally convex, and the Lie functor Lie is well-defined. As we will see, every finite-dimensional Lie algebra arises as the Lie algebra of some Lie group. In the general locally convex case this turns out to be a very subtle issue.

Example 13.10.2 (Groups of maps, Loop groups). Let M be a smooth manifold and H a Lie group (both potentially infinite-dimensional) with Lie algebra \mathfrak{h} . Then we obtain a natural topology on the group $G = C^\infty(M, H)$ as follows. The tangent bundle TH is a Lie group because the derivative of the smooth multiplication map on H defines a smooth multiplication map on TH . By iterating we get Lie group structures on all higher bundles $T^n H$. For each n we thus obtain a topological group $C(T^n M, T^n H)$ by using the topology of uniform convergence on compact subsets of $T^n M$, which coincides with the compact-open topology. We also observe that for two smooth maps $f_1, f_2 : M \rightarrow H$ the functoriality of T yields $T(f_1 \cdot f_2) = Tf_1 \cdot Tf_2$. Therefore the canonical map

$$C^\infty(M, H) \hookrightarrow \prod_n C(T^n M, T^n H), \quad f \mapsto (T^n f)_n \quad (13.10.2)$$

is a group homomorphism, so that the inverse image of the product topology on the right hand side is a group topology on $C^\infty(M, H)$. Therefore $C^\infty(M, H)$ always carries a natural structure of a topological group, even if M and H are infinite-dimensional.

Now we assume M is compact. Then these topological groups can be turned into Lie groups modeled on the space $\mathfrak{g} = C^\infty(M, \mathfrak{h})$. The charts of G are obtained from those of H as follows. If $\varphi_H : U_H \rightarrow \mathfrak{h}$ is a chart of H , then the set $U_G = \{f \in G : f(M) \subset U_H\}$ is an open subset of G . Assume, in addition, that $e \in U_H$ and $\varphi_H(e) = 0$. Then we use the map

$$\varphi_G : U_G \rightarrow \mathfrak{g}, \quad f \mapsto \varphi_H \circ f \quad (13.10.3)$$

is a chart of a e -neighborhood of G , and by combining it with left translations, we obtain an atlas of G defining a Lie group structure.

To calculate the Lie algebra of this group, we observe that for $m \in M$ we have the multiplication in local coordinates

$$(f *_G g)(m) = f(m) *_H g(m), \quad (13.10.4)$$

which can be expanded as usual in terms of the bilinear form $b(X, Y)$ whose antisymmetric part gives the Lie bracket

$$[f, g](m) = [f(m), g(m)]. \quad (13.10.5)$$

Thus $\text{Lie } C^\infty(M, H) = C^\infty(M, \mathfrak{h})$, endowed with the pointwise Lie bracket.

Remark 13.10.3. If M is a non-compact finite-dimensional manifold, one cannot expect $C^\infty(M, H)$ to be a Lie group. A typical example arises for $M = \mathbb{N}$ (a 0-dimensional manifold) and $H = U_1 = \mathbb{S}^1$. Then $C^\infty(M, H) \cong \mathbb{T}^{\mathbb{N}}$ is a topological group for which no e -neighborhood is contractible, so that it carries no smooth manifold structure.

Like with finite-dimensional Lie groups, we also find that the adjoint representation is smooth and that its derivative ad is the Lie bracket. Now we can move to Lie group actions on manifolds. If $\Phi : G \times M \rightarrow M$ is a smooth action, then we find that $T\Phi : TG \times TM \rightarrow TM$ is a smooth action of TG on TM and moreover the assignment of Killing vector fields $X \mapsto X_*$ is an (anti-)homomorphism of Lie algebras.

Example 13.10.4 (Diffeomorphism groups). Let M be a compact (hence also finite-dimensional) manifold and $\mathfrak{g} = \mathfrak{X}(M)$ the Lie algebra of smooth vector fields on M . We now explain how the group $\text{Diff}(M)$ can be turned into a Lie group modeled on \mathfrak{g} .

As we will see below, although $\text{Diff}(M)$ has a smooth exponential function, it is not a local diffeomorphism of a 0-neighborhood in \mathfrak{g} onto a e -neighborhood in G . Therefore we cannot use it to define charts for G . But there is an easy workaround. Let \mathfrak{g} be a Riemannian metric on M and $\exp : TM \rightarrow M$ be its exponential function, which assigns to $v \in T_m M$ the point $\gamma(1)$ where $\gamma : [0, 1] \rightarrow M$ is the geodesic segment with $\gamma(0) = m$ and $\dot{\gamma}(0) = v$. We then obtain a smooth map

$$\Phi : TM \rightarrow M \times M, \quad v \mapsto (m, \exp v), \quad v \in T_m M. \quad (13.10.6)$$

There exists an open neighborhood $U \subset TM$ of the zero section which Φ maps diffeomorphically onto an open neighborhood of the diagonal in $M \times M$. Now

$$U_{\mathfrak{g}} := \{X \in \mathfrak{X}(M) : X(M) \subset U\} \quad (13.10.7)$$

is an open subset of the Fréchet space $\mathfrak{X}(M)$, and we define the map

$$\varphi : U_{\mathfrak{g}} \rightarrow C^\infty(M, M), \quad \varphi(X)(m) := \exp X_m. \quad (13.10.8)$$

It is clear that $\varphi(0) = \text{id}_M$. One can show that after shrinking $U_{\mathfrak{g}}$ to a sufficiently small 0-neighborhood in the C^1 -topology on $\mathfrak{X}(M)$, we may achieve that $\varphi(U_{\mathfrak{g}}) \subset \text{Diff}(M)$. To see that $\text{Diff}(M)$ carries a Lie group structure for which φ is a chart, one has to verify that the group operations are smooth in a 0-neighborhood when transferred to $U_{\mathfrak{g}}$ via φ , so that

the standard method gives us a unique Lie group structure via left translations of this chart (see [40, Thm. 9.4.4]).

From the smoothness of the map $U_{\mathfrak{g}} \times M \rightarrow M$, $(X, m) \mapsto \varphi(X)(m) = \exp X_m$ it follows that the canonical left action $\Psi : \text{Diff}(M) \times M \rightarrow M$, $(\varphi, m) \mapsto \varphi(m)$ is smooth on an ϵ -neighborhood in $\text{Diff}(M)$, and hence smooth, because it is an action by smooth maps. The corresponding homomorphism of Lie algebras $\text{Lie Diff}(M) \rightarrow \mathfrak{X}(M)$ is given by

$$\dot{\Psi}(X)(m) = -T\Psi(X, 0) = -T_0 \exp|_{T_m M}(X_m) = -X_m, \quad (13.10.9)$$

i.e. $\dot{\Psi} = -\text{id}_{\mathfrak{X}(M)}$. This leads to

$$\text{Lie}(\text{Diff}(M)) = (\mathfrak{X}(M), -[\cdot, \cdot]). \quad (13.10.10)$$

The difference in sign is caused by the fact that we consider $\text{Diff}(M)$ as acting on M from the left. The tangent bundle of $\text{Diff}(M)$ can be identified with the set

$$T \text{Diff}(M) = \{X \in C^\infty(M, TM) : \pi_{TM} \circ X \in \text{Diff}(M)\}, \quad (13.10.11)$$

where π_{TM} is the tangent bundle projection. The multiplication in the group $T \text{Diff}(M)$ is given by the formula

$$X \times Y = \pi_{T^2 M} \circ TX \circ Y, \quad (13.10.12)$$

where $\pi_{T^2 M} : T^2 M \rightarrow TM$ is the bundle projection. Note that $\pi_{TM} \circ (X \times Y) = \pi_{TM} \circ X \circ \pi_{TM} \circ Y$, so π is a group homomorphism. Identifying $\varphi \in \text{Diff}(M)$ with the origin in $T_\varphi \text{Diff}(M)$, we get

$$X \cdot \varphi = \pi_{T^2 M} \circ TX \circ \varphi = X \circ \varphi, \quad \varphi \cdot X = \pi_{T^2 M} \circ T\varphi \circ X = T\varphi \circ X. \quad (13.10.13)$$

In particular, the adjoint representation on $T_e \text{Diff}(M) = \mathfrak{X}(M)$ is given by push-forwards,

$$\text{Ad}_\varphi X = T\varphi \circ X \circ \varphi^{-1} = \varphi_* X, \quad (13.10.14)$$

according to (8.3.21).

We now briefly review how other structural results extend to the infinite-dimensional case. As alluded to above, the exponential map is not necessarily locally diffeomorphic. For this purpose we define a special class of Lie groups.

Definition 13.10.5 (Regular Lie group). A Lie group G is called regular if for each $\xi \in C^\infty([0, 1], \mathfrak{g})$ the initial value problem

$$\dot{\gamma}(t) = L_{\gamma(t)*}\xi(t), \quad \gamma(0) = e \quad (13.10.15)$$

has a solution $\gamma_\xi \in C^\infty([0, 1], G)$ and the evolution map

$$\text{evol}_G : C^\infty([0, 1], \mathfrak{g}) \rightarrow G, \quad \xi \mapsto \gamma_\xi(1) \quad (13.10.16)$$

is smooth. For a regular Lie group G we define the exponential map

$$\exp : \mathfrak{g} \rightarrow G, \quad \exp X = \gamma_X(1) = \text{evol}_G(x), \quad (13.10.17)$$

where $X \in \mathfrak{g}$ is treated as the constant function $[0, 1] \rightarrow \mathfrak{g}$.

Remark 13.10.6. (a) As a direct consequence of the existence of solutions to ODEs on open domains of Banach spaces and their smooth dependence on parameters, every Banach-Lie group is regular.

- (b) Let A be a unital Banach algebra and A^\times its unit group (group of invertible elements). Since A is a CIA (*continuous inverse algebra*: Hausdorff locally convex unital algebra such that A^\times is open and inversion map $A^\times \rightarrow A$ is continuous), A^\times is a Lie group. For $x \in A$ the corresponding left-invariant vector field on A^\times is given by $x_L(a) = ax$, and the unique solutions of the above initial value problem are given by $\gamma(t) = \exp tx$ given by the convergent exponential Taylor series. This implies that \exp_A is a smooth exponential function of the Lie group A . This remains true for each Mackey complete CIA: the exponential series converges and \exp_A defines a smooth exponential map on A^\times .
- (c) Although it might be hard to verify it in concrete situations, all “known” Lie groups modeled on Mackey complete spaces are regular. For example we do not know if all unit groups of Mackey complete CIAs are regular, but we have just seen in (b) that they always have a smooth exponential function. If the model space is not Mackey complete, one can easily obtain non-regular groups. E.g. the unit Lie group of the algebra of rational functions on $[0, 1]$ is not regular because e^f is not rational for a rational f .
- (d) A locally convex space V is a Lie group under the addition operation iff it is Mackey complete. Regularity is then inherited by all abelian Lie groups V/Γ where Γ is a discrete subgroup of V .
- (e) If H is a Lie group with a smooth exponential map $\exp_H : \mathfrak{h} \rightarrow H$ and M is a compact manifold, then we obtain an exponential map of the group $G = C^\infty(M, H)$ by

$$\exp_G : \mathfrak{g} = C^\infty(M, \mathfrak{h}) \rightarrow G, \quad \xi \mapsto \exp_H \circ \xi. \quad (13.10.18)$$

An important theorem states that given an exact sequence of Lie groups $0 \rightarrow N \rightarrow \hat{G} \rightarrow G \rightarrow 0$ (so \hat{G} is a principal N -bundle), \hat{G} is regular iff G and N are regular.

Initial value problems on regular Lie groups of the form

$$\dot{\eta}(t) = [\eta(t), \xi(t)], \quad \eta(0) = X \quad (13.10.19)$$

have a unique solution given by $\eta(t) = \text{Ad}_{\xi(t)}^{-1} X$ just like in finite-dimensional Lie groups.

Remark 13.10.7. Suppose two functions $\xi, \eta : I^2 \rightarrow \mathfrak{g}$ (with I^2 parametrized by (t, s)) satisfy the Maurer-Cartan equation in the form

$$\partial_s \xi - \partial_t \eta = [\xi, \eta]. \quad (13.10.20)$$

Then we define a smooth function $f : I^2 \rightarrow G$ by

$$f(t, 0) = \gamma_{\xi(\cdot, 0)}(t), \quad f(t, s) = f(t, 0) \cdot \gamma_{\eta(t, \cdot)}(s). \quad (13.10.21)$$

Since the map $I \rightarrow C^\infty(I, \mathfrak{g})$, $t \mapsto \eta(t, \cdot)$ is smooth, f is a smooth function. We have

$$L_{f^*}^{-1} f = \hat{\xi} \cdot dt + \eta \cdot ds, \quad \hat{\xi}(t, 0) = \xi(t, 0). \quad (13.10.22)$$

The Maurer-Cartan equation for $L_{f^*}^{-1} f$ reads $\partial_s \hat{\xi} - \partial_t \eta = [\hat{\xi}, \eta]$, so subtracting it from the original equation we get

$$\partial_s (\xi - \hat{\xi}) = [\xi - \hat{\xi}, \eta]. \quad (13.10.23)$$

Since $(\xi - \hat{\xi})(t, 0) = 0$, the uniqueness of solutions applied to $\eta(t, \cdot)$ implies that $\xi = \hat{\xi}$ on all of I^2 . Thus $L_{f^*}^{-1} f = \xi dt + \eta ds$.

The following familiar theorem holds for solutions of the Maurer-Cartan equation on regular Lie groups. We call a one-form $\alpha \in \Omega^1(M, \mathfrak{g})$ integrable if there exists a smooth function $f : M \rightarrow G$ such that $L_f^* df = \alpha$.

Theorem 13.10.8 (Fundamental Theorem for Lie group valued functions). *Let M be a simply connected manifold and G a regular Lie group. Then $\alpha \in \Omega^1(M, \mathfrak{g})$ is integrable iff it satisfies the Maurer Cartan equation*

$$d\alpha + \frac{1}{2} [\alpha, \alpha] = 0. \quad (13.10.24)$$

Using this, one then shows, as we did in the finite-dimensional case, that every homomorphism between the Lie algebras of two simply connected Lie groups generates a unique Lie group homomorphism.

Definition 13.10.9. A Lie group G is called locally exponential if it has a smooth exponential map $\exp : \mathfrak{g} \rightarrow G$ which is diffeomorphic on some neighborhood of $0 \in \mathfrak{g}$.

Under the assumption that G is locally exponential, we can recover many of the analytic results that we've shown in the finite-dimensional case: every continuous homomorphism

of such groups is actually smooth, and Lie algebra homomorphisms $\mathfrak{g} \rightarrow \mathfrak{h}$ with G simply connected generate unique Lie group homomorphisms.

Example 13.10.10 ($\text{Diff}(\mathbb{S}^1)$ is not locally exponential). We will show that every ϵ -neighborhood in $\text{Diff}(\mathbb{S}^1)$ contains elements which do not lie in a OPS. Let $G = \text{Diff}_+(\mathbb{S}^1)$ be the identity component of $\text{Diff}(\mathbb{S}^1)$. Its universal covering group \tilde{G} is, as we will show,

$$\tilde{G} = \{\varphi : \text{Diff}(\mathbb{R}) : \forall x \in \mathbb{R} \quad \varphi(x + 2\pi) = \varphi(x), \varphi'(x) > 0\}. \quad (13.10.25)$$

The covering map π is the composition $\varphi \mapsto q \circ \varphi \circ q^{-1}$ with the quotient map $q : \mathbb{R} \rightarrow \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$ (this is well-defined by periodicity of φ). The Lie group structure on \tilde{G} can be defined by a single global chart. Denoting by $C_{2\pi}^\infty(\mathbb{R}, \mathbb{R})$ the Fréchet space of 2π -preperiodic smooth functions on \mathbb{R} , which is a closed subspace of $C^\infty(\mathbb{R}, \mathbb{R})$, we have the open convex subset

$$U = \{\varphi \in C_{2\pi}^\infty(\mathbb{R}, \mathbb{R}) : \varphi' > -1\} \quad (13.10.26)$$

and the bijective map

$$\Phi : U \rightarrow \tilde{G}, \quad \Phi(f)(x) = x + f(x). \quad (13.10.27)$$

It's easy to see that $\Phi(f)(x + 2\pi) = \Phi(f)(x) + 2\pi$ and $\Phi(f)' > 0$. The group operations in \tilde{G} are defined via this chart by

$$m(f, g)(x) = f(g(x) + x) - x, \quad i(f)(x) = (f + \text{id}_{\mathbb{R}})^{-1}(x) - x, \quad (13.10.28)$$

which can be shown directly to be smooth. We thus obtain on \tilde{G} the structure of a Lie group such that Φ is a diffeomorphism. In particular \tilde{G} is contractible and therefore simply connected. This proves that \tilde{G} is the universal covering group of G .

Now we construct certain elements of \tilde{G} which are close to the identity. For $0 < \epsilon < \frac{1}{n}$ consider the function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x + \frac{\pi}{n} + \epsilon \sin^2(nx) \quad (13.10.29)$$

and observe that $f \in \tilde{G}$ follows from $f'(x) = 1 + 2\epsilon n \sin(nx) \cos(nx) = 1 + \epsilon n \sin(2nx) > 0$. For large n fixed and $\epsilon \rightarrow 0$ we get elements in \tilde{G} which are arbitrarily close to $\text{id}_{\mathbb{R}}$.

$\pi(f)$ has a unique periodic orbit of order $2n$ on \mathbb{S}^1 : under $\pi(f)$ the point $q(0) \in \mathbb{S}^1$ is mapped to π/n etc., so we obtain the orbit $q(0) \rightarrow q(\pi/n) \rightarrow \dots \rightarrow q((2n-1)\pi/n) \rightarrow q(0)$. For $0 < x_0 < \pi/n$ we have for $x_1 = f(x_0)$:

$$x_0 + \frac{\pi}{n} < x_1 < \frac{2\pi}{n}, \quad (13.10.30)$$

and for $x_n = f(x_{n-1})$ the relations

$$0 < x_0 < x_1 - \frac{\pi}{n} < x_2 - \frac{2\pi}{n} < \dots < \frac{\pi}{n}. \quad (13.10.31)$$

Therefore $x_k - x_0 \in 2\pi\mathbb{Z}$ for each $k \in \mathbb{N}$, and hence the orbit of $q(x_0)$ under π^*f is not finite. This proves that $\pi(f)$ has a unique periodic orbit and its order is $2n$. Now we prove that $\pi(f) \neq g^2$ for all $g \in \text{Diff}(\mathbb{S}^1)$. Every periodic orbit of g is a periodic orbit of g^2 and vice versa. If the period of x under g is odd, then its period under g^2 is the same. If the period of x is $2m$, then its orbit under g breaks up into two orbits under g^2 , each of order m . Therefore g^2 can never have a single periodic orbit of even order, and this proves that $\pi(f)$ has no square root in $\text{Diff}(\mathbb{S}^1)$. It follows that $\pi(f)$ doesn't belong to any OPS.

- Remark 13.10.11.** (a) If M is compact, then one can show that the identity component $\text{Diff}(M)_0$ is a simple group. Being normal in $\text{Diff}(M)_0$, the subgroup $\langle \exp \mathfrak{X}(M) \rangle$ coincides with $\text{Diff}(M)_0$. Hence every diffeomorphism homotopic to the identity is a finite product of exponentials.
- (b) Although $\text{Diff}(M)_0$ is a simple Lie group, its Lie algebra $f\mathcal{X}(M)$ is far from being simple. For each subset $K \subset M$ the set $\mathfrak{X}_K(M)$ of all vector fields supported in K is a Lie algebra ideal which is proper if K is not dense.

Finally, we note that the correspondence between analytic (immersed) Lie subgroups and Lie subalgebras holds for locally exponential groups. This includes the Analytic Subgroup Theorem. However, the Lie's Third Fundamental Theorem does not hold. Lie algebras of vector fields provide a counterexample:

Theorem 13.10.12 (Lempert). *If M is a compact manifold, then the Fréchet-Lie algebra $\mathfrak{X}(M)_{\mathbb{C}}$ is not integrable to a regular Lie group.*

Moreover, $\mathfrak{X}(M)$ for non-compact M is not integrable at all.

Part III

Structured Geometry

14 Principal and Associated Bundles

14.1 Principal fiber bundles

Recall that in Definition 13.7.6 we defined principal bundles as free and proper Lie group actions, and provided an equivalent characterization in terms of equivariant local trivializations. We now make the latter definition explicit.

Definition 14.1.1 (Principal fiber bundle II). Let (P, G, Φ) be a free Lie group action, let M be a manifold and let $\pi : P \rightarrow M$ be a smooth mapping. The tuple $(P \xrightarrow{\pi} M, G, \Phi)$ is called a principal G -bundle if for every $m \in M$ there exists an open neighborhood U_p of m and a diffeomorphism $\chi : \pi^{-1}(U_p) \rightarrow U_p \times G$ such that

1. χ intertwines Φ with the G -action on $U \times G$ by translations of the second component (left if Φ is left, resp. right if it is right).
2. $\text{pr}_{U_p} \circ \chi = \pi$ on $\pi^{-1}(U_p)$.

For brevity we will often denote principal bundles by $P(G, M)$.

The existence of local trivializations implies that π is a smooth submersion. The fibers of a surjective submersion are embedded submanifolds, all diffeomorphic to the structure group G . They are also the orbits of the G -action on P . Following convention, we will usually assume that Φ is a right action and call P a *right* principal G -bundle (not to be confused with the left G -action on the typical fiber).

Remark 14.1.2.

1. Principal bundles can also be understood in terms of G -structures. Namely a G -bundle $(P \xrightarrow{\pi} M, G \curvearrowright F, \mathcal{G})$ is principal iff the action $G \curvearrowright F$ is free and transitive.
2. Recall that a G -invariant structure can be naturally lifted from the typical fiber to the fibers of the G -bundle itself. However, the Lie group structure of G is not translation-invariant: e.g. the identity element is not invariant under translations. Therefore *the fibers of a PFB do not carry a natural group structure*. In particular, unlike **VBS**, they have no distinguished elements and therefore no analog of a zero section. Instead, the fibers are merely principal homogeneous spaces of G (that is, $G/\{e\}$).
3. Another feature that distinguishes **PFBs** from all other G -bundles is that the typical fiber G carries *two* commuting actions of G . The bundle construction via transition functions has those functions act on G , say, from the left, which precludes the existence of any left

G -action on the bundle itself (unless G is abelian). However the action of G on itself by right translations is left untouched by the bundle construction, and this is the action with respect to which the trivializations are equivariant in the above definition. Even if G is abelian, then the two actions coincide anyway, so there still won't be two actions on the bundle.

Since all of these facts are easier to state explicitly in the definition of a principal bundle, Definition 14.1.1 is the most common one in the literature.

- Example 14.1.3.**
1. $\text{pr}_M : M \times G \rightarrow M$ with the right action of G on the second factor is a trivial G -PFB.
 2. The direct product $P_1 \times P_2$ of two PFBs with respective structure groups G_1, G_2 over bases M_1 and M_2 , has a natural structure of a principal $G_1 \times G_2$ -bundle over $M_1 \times M_2$.
 3. If $H \triangleleft G$ is a closed subgroup, then H acts on G freely by right translation and this action is proper. Thus $G \rightarrow G/H$ is a principal H -bundle.

Proposition 14.1.4 ([17, Prop. 1.1.6]). *Local trivializations of a PFB are in bijective correspondence with its local sections. In particular, a PFB is trivial iff it admits a global section.*

Proof. If $\xi : P \rightarrow M \times G$ is a global trivialization, then we set $s(m) = \xi^{-1}(m, e)$, where $e \in G$ is the identity. This is a smooth global section of P . Conversely, given a global section $s : M \rightarrow P$, for every point $p \in P$ there exists a unique group element $g(p) \in G$ such that $p = \Phi_{g(p)}(s(\pi(p)))$. This defines an equivariant smooth map $g : P \rightarrow G$, so that $g(\Phi_h(p)) = g(p)h$. Thus $(M, \pi \times g)$ is a global trivialization.

The statement for local sections follows by considering trivial restrictions of the bundle of the form $P|_U$. ■

Now we need to specify the notion of morphism of PFBs. Recall that a morphism between G_1 - and G_2 -structures is a bundle morphism whose fiberwise representatives lie in $G_1 \times G_2$ -orbits of some canonically chosen maps between the typical fibers. Since the typical fibers of principal bundles are the Lie groups themselves, it is clear that these canonical maps should be chosen as Lie group homomorphisms. The $G_1 \times G_2$ -orbits of homomorphisms are exactly equivariant maps, as we now show.

Definition 14.1.5 (λ -equivariant map). Let $G_1 \curvearrowright F_1$ and $G_2 \curvearrowright F_2$ be two right Lie group actions and let $\lambda : G_1 \rightarrow G_2$ be a Lie group homomorphism. A smooth map $\varphi : F_1 \rightarrow F_2$ is called (right) λ -equivariant if $\varphi(f \cdot g_1) = \varphi(f) \cdot \lambda(g_1)$ for all $f \in F, g_1 \in G_1$.

Proposition 14.1.6. *Let $\lambda : G \rightarrow G'$ be a Lie group homomorphism and let G, G' act on themselves by right translations. Then a smooth map $\varphi : G \rightarrow G'$ is right λ -equivariant iff it has the form $\varphi(g) = g' \lambda(g)$ for some $g' \in G'$.*

Proof. Let φ be right λ -equivariant. Then $\varphi(g) = \varphi(e \cdot g) = \varphi(e)\lambda(g)$, so setting $g' = \varphi(e)$ proves the assertion. The other direction is obvious. ■

Recall that the maps $\Phi_{\alpha a} : U_{\alpha a} \times (G \times G') \rightarrow C^\infty(G, G')$ defined in (7.4.7) are equivariant with respect to the $G \times G'$ -action on $C^\infty(G, G')$ so that requiring their values to be in an orbit of a homomorphism λ means exactly requiring that these values take the form $h \mapsto h'\lambda(h)$ (strictly speaking it should be $h'\lambda(\tilde{h}^{-1}h)$ but we can just redefine h' since λ is a homomorphism). Note that h' can be a function of p and the indices α, a . Thus the definition of a LieGr-bundle simply requires that the maps $\Phi_{\alpha a}$ take values in right λ -equivariant maps $G \rightarrow G'$. We thus arrive at the following definition of morphisms of PFBs and the final alternative definition of principal bundles as LieGr-bundles.

Definition 14.1.7 (Morphism of principal bundles). A morphism between two (right) PFBs $(P \xrightarrow{\pi} M, G, \Phi)$ and $(P' \xrightarrow{\pi'} M', G', \Phi')$ consists of a Lie group homomorphism $\lambda : G \rightarrow G'$ and a (right) λ -equivariant bundle morphism $\vartheta : P \rightarrow P'$, i.e. such that for all $g \in G$

$$\vartheta \circ \Phi_g = \Phi'_{\lambda(g)} \circ \vartheta. \quad (14.1.1)$$

Definition 14.1.8 (Principal fiber bundle III). A (right) PFB is a LieGr-bundle in the sense of Definition 7.4.1, using the action of every Lie group on itself by (left) translations. The resulting category LieGr-Bun^∞ is denoted Prin .

Upon restricting to a specific Lie group G and morphisms with $\lambda = \text{id}_G$ one gets the category Prin^G , which coincides with the category Bun^L (Definition 7.4.2) where L is the action of G on itself by left translations. Upon further restricting the base to a specific manifold M and morphisms to those covering the identity one gets Prin_M^G (all morphisms in this category are isomorphisms).

It might seem that Definition 14.1.1 contains more structure in the form of *equivariant* trivializations, however the bundle construction of a LieGr-bundle out of a cocycle (say acting on the left on the typical fiber G) allows us to *lift* the right G -action to the whole bundle. Right G -equivariance of trivializations is therefore automatic under this definition.

Remark 14.1.9 (Special morphisms, bundle reduction).

1. As in all bundles, ϑ maps fibers to fibers, thus it induces a map $\theta : M \rightarrow M'$ which we will call *the horizontal part* of ϑ . It is smooth by local triviality and if (ϑ, λ) is an isomorphism the it is a diffeomorphism.
2. We call the morphism vertical if $M = M'$ and $\theta = \text{id}_M$. If $G = G'$ and $\lambda = \text{id}_G$, then ϑ is called a G -morphism. By local triviality, every vertical G -morphism is a diffeomorphism and hence an isomorphism.
3. If θ and λ are injective immersions, then ϑ is one as well. In this case P is called a

subbundle of P' and we write $P < P'$.

- (a) If, additionally, θ and λ are embeddings, then we write $P \lhd P'$. In the latter case P may be identified with the image of the morphism ϑ in P' .
- (b) If, additionally, $M = M'$ and $\theta = \text{id}_M$, then P is called a λ -reduction of P' . In this case one says that G is a reduction of the structure group G' . Two reductions are said to be equivalent if they differ by a vertical automorphism of P .

Remark 14.1.10. 1. Let $\vartheta : P \rightarrow P'$ be a principal G -bundle morphism covering $\theta : M \rightarrow M'$. The induced mapping

$$P \rightarrow \theta^*P', \quad p \mapsto (\pi(p), \vartheta(p)), \quad (14.1.2)$$

is a vertical isomorphism and ϑ decomposes into the composition of this isomorphism with the natural principal G -bundle morphism $\theta^*P' \rightarrow P'$.

- 2. The following is an important class of pullback bundles. Let $P(M, G)$ and $P'(M, G')$ be principal bundles. Then $P \times P'$ is a principal $(G \times G')$ -bundle over $M \times M$. Let $\Delta : M \rightarrow M \times M$ be the diagonal embedding. Then $\Delta^*(P \times P')$ is a principal $(G \times G')$ -bundle over M . We shall denote it by $P \times_M P'$ and call it *the fibered product of P and P'* , also called a *spliced product* in physics.

14.2 Associated Bundles

In our investigation of the consequences of the general bundle construction in the case of principal bundles it remains to examine what the compatibility condition (7.4.8) means for morphisms of principal bundles. According to the above discussion, $\Phi_{\alpha a}$'s can be written as

$$\Phi_{\alpha a}(p, h, h')(g) = h' \hat{g}'_{\alpha a}(p) \lambda(h^{-1} g), \quad (14.2.1)$$

where the functions $\hat{g}'_{\alpha a} : U_{\alpha a} \rightarrow G'$ are uniquely determined by this equation. Together with the homomorphism λ , these functions completely describe the bundle morphism. Denoting by $\theta : M \rightarrow M'$ the horizontal part of the bundle morphism ϑ , the transformation law (7.4.8) becomes

$\hat{g}'_{\beta b}(m) = t'_{ba}(\theta(m)) \hat{g}'_{\alpha a}(m) \lambda(t_{\beta \alpha}(m))^{-1}.$

(14.2.2)

This formula suggests viewing the functions $\{\hat{g}'_{\alpha a}\}_{\alpha, a}$ as local representatives of a global section of some bundle with typical fiber G' whose transition functions are $(\lambda \circ t_{\beta \alpha}, t'_{ba}(\theta(m)))$ acting on G' as in the formula above. To understand what this bundle is we will need the concept of associated bundles.

Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a right principal bundle and let (F, G, σ) be a left Lie group action.

Let $\check{\sigma}$ be the right action associated with σ :

$$\check{\sigma} : F \times G \rightarrow F, \quad (f, g) \mapsto \check{\sigma}_g(f) := \sigma_{g^{-1}}(f). \quad (14.2.3)$$

σ can also be a right action, in which case we set $\check{\sigma} = \sigma$. Since the G -action Φ is free, the direct product action $\Phi \times \check{\sigma}$ is free as well. By Remark 13.5.7, it is proper. Thus, by Corollary 13.7.1, the orbit space $(P \times F)/G$ inherits a unique smooth structure. Since the natural projection $P \times F \rightarrow P$ is equivariant, it induces a smooth surjective mapping

$$\pi_F : (P \times F)/G \rightarrow P/G = M, \quad \pi_F([(p, f)]) = \pi(p). \quad (14.2.4)$$

Local triviality of P implies local triviality of this map, so it turns $(P \times F)/G \xrightarrow{\pi_F} M$ into a fiber bundle. Indeed, recall from Proposition 14.1.4 that local trivializations of P correspond to local sections. Let $s : U \rightarrow P$ be a local section of P . Then the map

$$U \times F \rightarrow \pi_F^{-1}(U), \quad (m, f) \mapsto [(s(m), f)] \quad (14.2.5)$$

is a diffeomorphism projecting to the identity on U . The inverse map $\xi : \pi_F^{-1}(U) \rightarrow U \times F$ yields a local trivialization of $P \times_G F$.

Definition 14.2.1 (Twisted product, associated bundle). If $(P \xrightarrow{\pi} M, G, \Phi)$ is a principal bundle and (F, G, σ) a G -manifold, then the orbit manifold defined above is called the twisted product of P and F and is denoted by

$$P \times_G F := (P \times F)/G. \quad (14.2.6)$$

The map $P \times_G F \xrightarrow{\pi_F} M$ provides it with a natural structure of a fiber bundle. Such a bundle is said to be associated to the principal bundle P and the G -manifold F .

More generally the quotient of the product of a right G -space and a left G -space as above is called a *balanced product*. In the smooth case, for these products to have well-defined smooth structures, we require the group actions to be free and proper, so that orbit manifolds are well defined.

Example 14.2.2.

1. If X is a right G -manifold and G acts on itself by left translations, then $X \times_G G \cong X$.
2. We say that Y is a (G, H) -manifold if it carries a left action of G and a right action of H , and these two actions commute. Then balanced products are associative up to isomorphism in the sense that if X is a right G -manifold, Y is a (G, H) -manifold, and Z is a left H -manifold, then there is a natural diffeomorphism

$$(X \times_G Y) \times_H Z \cong X \times_G (Y \times_H Z). \quad (14.2.7)$$

3. If $H \triangleleft G$ is a closed subgroup, then G can be viewed as a (G, H) -manifold. Combining the previous two examples, we conclude that the symbol $\times_G G$ (or $G \times_G$) can be cancelled whenever it occurs:

$$X \times_G G \times_H Y \cong X \times_H Y. \quad (14.2.8)$$

4. Taking Y to be a point in the last example, we have for a right G -manifold X and a closed subgroup $H \triangleleft G$ a natural diffeomorphism

$$X \times_G (G/H) \cong X/H. \quad (14.2.9)$$

In particular, if P is a principal G -bundle, then we can identify P/H with the total space of the associated bundle $P \times_G (G/H)$.

The following proposition is an exercise.

Proposition 14.2.3 ([17, Prop. 1.2.2]). *Let $P(M, G)$ and $P'(M', G')$ be principal bundles and let (F, G, σ) and (F', G', σ') be left Lie group actions. Let (ϑ, λ) be a principal bundle morphism $P \rightarrow P'$ covering $\theta : M \rightarrow M'$ and let $T : F \rightarrow F'$ be an equivariant map (morphism of Lie group actions). Then $\vartheta \times T : P \times F \rightarrow P' \times F'$ induces a bundle morphism $P \times_G F \rightarrow P' \times_{G'} F'$ covering θ .*

This means that we have constructed an “association functor” which takes a principal G -bundle P and a G -action σ and produces the associated bundle. We will denote this functor on objects by

$$[] : \text{Prin}^G \times G\text{-Man}^\infty \rightarrow \text{Bun}^\infty, \quad (P, \sigma) \mapsto P^{[\sigma]} := P \times_G F. \quad (14.2.10)$$

If σ is a faithful action, then this restricts to a functor $[\sigma] : \text{Prin}^G \rightarrow \text{Bun}^\sigma$ (see Definition 7.4.2). We will show below that $[\sigma]$ is an equivalence of categories.

Remark 14.2.4. 1. The G -action σ on F need not be faithful. When it isn’t, this simply means that the structure group of $P \times_G F$ is not G , but the image of G under some homomorphism. This subtlety is important in applications where the associated bundle is constructed using non-faithful representations. For example, spinor bundles are finite-rank VBs associated to principal Spin-bundles. Since most spin groups are either nonlinear or their low-dimensional representations are all non-faithful (the smallest faithful representation of Spin_{2n} is 2^n -dimensional), the actual structure group of a spinor bundle won’t be a spin group.
2. The local trivialization $\xi : \pi_F^{-1}(U) \rightarrow U \times F$ constructed above for $P \times_G F$ can be verified to take the form

$$\xi([p, f]) = (\pi_F([p, f]), \sigma_{\varphi(p)}(f)), \quad (14.2.11)$$

where $\varphi = \text{pr}_G \circ \xi : U \rightarrow G$ is the projection of the local trivialization $\xi : \pi^{-1}(U) \rightarrow U \times G$ of P onto the group factor.

3. The natural quotient map $\iota : P \times F \rightarrow P \times_G F$ induces for every $p \in P$ a map

$$\iota_p : F \rightarrow P \times_G F, \quad \iota_p(f) = [(p, f)], \quad (14.2.12)$$

whose image is contained in the fiber over $\pi(p)$. Moreover ι_p is equivariant,

$$\iota_{\Phi_g(p)} = \iota_p \circ \sigma_g. \quad (14.2.13)$$

Therefore $\iota_p : F \rightarrow \pi_F^{-1}(\pi(p))$ is bijective. From (14.2.11) we read off

$$\text{pr}_2 \circ \xi \circ \iota_p = \sigma_{\varphi(p)}, \quad (14.2.14)$$

for any local trivialization (U, ξ) such that $\pi(p) \in U$. Since this is a diffeomorphism of F , we conclude that i_p is a diffeomorphism. In other words, the typical fiber of $P \times_G F$ is F .

4. If $t_{\beta\alpha} : U_{\alpha\beta} \rightarrow G$ are the transition functions of P , then it is easy to see from (14.2.11) that they are also the transition functions of $P \times_G F$: choosing $f \in F$ and $p \in \pi^{-1}(m)$,

$$\xi_\beta \circ \xi_\alpha^{-1}(m, f) = \xi_\beta([(p, \sigma_{\varphi_\alpha(p)}^{-1}(f))]) = (m, \sigma_{\varphi_\beta(p)\varphi_\alpha(p)^{-1}}(f)) = (m, \sigma_{t_{\beta\alpha}(m)}(f)). \quad (14.2.15)$$

Therefore the easiest way to conceptualize the associated bundle $P \times_G F$ is as a bundle with the same transition functions as P but with the typical fiber replaced by F . This is why we usually assume that σ is a left action if Φ is a right one.

The following is an exercise.

Proposition 14.2.5. *There is a natural bundle isomorphism $f^*(P \times_G F) \cong (f^*P) \times_G F$. This can also be written as $N \times_M (P \times_G F) \cong (N \times_M P) \times_G F$, but it is important to remember the very different natures of the two products.*

Example 14.2.6. 1. Let $P(M, G)$ be a PFB and let $H \triangleleft G$ be a closed subgroup. Then by the Closed Subgroup Theorem 12.4.2, H is an embedded Lie subgroup, and we get the homogeneous space G/H with a G -action by left translation. Then $P \times_G G/H$ is an associated bundle over M with typical fiber being a transitive G -manifold. One can show the following:

- (a) As a bundle over M , the bundle $P \times_G G/H$ is isomorphic to the quotient P/H endowed with the natural bundle structure induced from P .
- (b) P may be viewed as a principal H -bundle over $P \times_G G/H$.

2. If $P(M, G)$ is a **PFB**, $E = P \times_G F$ is an associated bundle, and $f : N \rightarrow M$ a smooth map, then f^*E is naturally associated to f^*P via the isomorphism

$$f^*E \rightarrow f^*P \times_G F, \quad (y, [(p, f)]) \mapsto [((y, p), f)]. \quad (14.2.16)$$

3. If $E_i = P_i \times_{G_i} V_i, i = 1, 2$, are **VBS**s associated with $P_i(M, G_i)$ via representations σ_i , respectively, then we can build $(P_1 \times P_2) \times_{(G_1 \times G_2)} (V_1 \otimes V_2)$ over $M \times M$ using the product representation $\sigma_1 \otimes \sigma_2$. Then we can take the pullback along the diagonal map $\Delta : M \rightarrow M \times M$. Using point 2, we obtain

$$\Delta^*((P_1 \times P_2)_{G_1 \times G_2}(V_1 \otimes V_2)) = (P_1 \times_M P_2) \times_{(G_1 \times G_2)} (V_1 \otimes V_2). \quad (14.2.17)$$

It is easy to show that this bundle is isomorphic to the tensor product bundle $E_1 \otimes E_2$, that is, $E_1 \otimes E_2$ is naturally associated with the fiber product $P_1 \times_M P_2$.

Recall that all categories Bun^σ for different *faithful* G -actions σ are equivalent to each other. Therefore the category $\text{Prin}^G = \text{Bun}^L$ is equivalent to any of the categories Bun^σ of G -bundles. The functor providing this equivalence is $C_{\sigma L} : \text{Prin}^G \rightarrow \text{Bun}^\sigma$, and its pseudo-inverse is $C_{L\sigma}$.

Definition 14.2.7 (Frame bundle). If $(E \xrightarrow{\pi} M, G \curvearrowright^\sigma F, \mathcal{G})$ is a G -bundle, then the principal G -bundle

$$\mathcal{P}(E) := C_{L\sigma}(E), \quad (14.2.18)$$

uniquely defined by the maximal cocycle of E , is called the G -frame bundle of E . If E is a vector bundle, it is often denoted $\text{Fr}(E)$ or $\text{Fr}_G(E)$.

Since we've already established that P and $P^{[\sigma]}$ have the same transition functions, this means that there is a natural isomorphism between the functors $[\sigma]$ and $C_{\sigma L}$. We thus have

Proposition 14.2.8. *If $G \curvearrowright^\sigma F$ is a faithful Lie group action, then the association functor $[\sigma] : P \mapsto P^{[\sigma]}$ establishes an equivalence of categories Prin^G and Bun^σ . Its pseudoinverse \mathcal{P} reconstructs the principal G -frame bundle.*

Remark 14.2.9. The above equivalence is intuitive: with the action σ fixed, bundles on both sides of the equivalence are determined up to isomorphism by a G -valued cocycle. One crucial implication of this equivalence is:

To classify all fiber bundles it suffices to classify principal bundles.

An even deeper implication is the following. Our original definition of fiber bundles intrinsically relied on atlases of local trivializations and corresponding G -structures. Meanwhile principal bundles can be defined purely geometrically as free proper Lie group actions. The

association functor is also defined purely geometrically as a quotient of a Lie group action. We thus have

General fiber bundles can be *defined* as quotients of the form $(P \times F)/G$.

As we will see, this also leads to a coordinate-free definition of G -structures. This situation is fundamentally different from the definition of general manifolds, in which the references to local Euclidean structures cannot be circumvented.

Example 14.2.10 (Covering spaces). Recall that for a discrete (and countable) group G , principal G -bundles are nothing but regular covering maps whose group of deck transformations is G . In particular, the universal covering of M is regular with $G = \pi_1(M)$. Any other covering of M has structure group that is the image of $\pi_1(M)$ under some homomorphism $\lambda : \pi_1(M) \rightarrow \text{Diff}(F)$. Thus every covering can be written as $P^{[\lambda]}$ for some regular covering P .

An important subclass of associated bundles are associated principal bundles. If P is a principal G -bundle and $\lambda : G \rightarrow H$ is a Lie group homomorphism, then we can let G on H by left translations via λ , i.e. $(g, h) \mapsto \lambda(g)h$. The resulting associated bundle is denoted

$$P^{[\lambda]} := P \times_G H. \quad (14.2.19)$$

Since left and right translations on H commute, the action of H on $P \times H$ by right translations on the second factor descends to a free right action on $P^{[\lambda]}$, which turns $P^{[\lambda]}$ into a principal H -bundle over the same base. Its transition functions as an H -bundle are the compositions of λ with the transition functions of P . This operation has the following functorial properties (exercise):

Proposition 14.2.11 (Associated principal bundles [17, Prop. 1.2.5]). *Let $\lambda : G \rightarrow H$ be a Lie group homomorphism and let P, P' be principal G -bundles over, respectively, M and M' .*

1. *A morphism $\vartheta : P \rightarrow P'$ of principal G -bundles induces a morphism of associated principal H -bundles via*

$$P^{[\lambda]} \rightarrow P'^{[\lambda]}, \quad [(p, h)] \mapsto [(\vartheta(p), h)]. \quad (14.2.20)$$

2. *Any $f : N \rightarrow M$ smooth map induces a natural vertical isomorphism*

$$f^*(P^{[\lambda]}) \rightarrow (f^*P)^{[\lambda]}, \quad (m, [(p, h)]) \mapsto [(f(m), p, h)]. \quad (14.2.21)$$

3. *If $(P_i, G_i, \lambda_i, H_i)$ are two tuples as above, then the rearrangement $(P_1 \times P_2) \times (H_1 \times H_2) \rightarrow (P_1 \times H_1) \times (P_2 \times H_2)$ induces a vertical isomorphism*

$$(P_1 \times P_2)^{[\lambda_1 \times \lambda_2]} \cong P_1^{[\lambda_1]} \times P_2^{[\lambda_2]}. \quad (14.2.22)$$

Therefore in this special case we have a functor of the categories of principal bundles $[\lambda] : \text{Prin}^G \rightarrow \text{Prin}^H$. Since this functor “forgets” the action of G on H (in other words, the structure group of an H -bundle is always H), it is an equivalence of categories only if λ is an isomorphism, in which case its inverse is $[\lambda^{-1}]$.

Example 14.2.12. 1. If V is a \mathbb{K} -vector space of dimension k , then the space $\text{Fr}(V)$ of linear isomorphisms (frames) $V \rightarrow \mathbb{K}^k$ is a principal homogeneous $\text{GL}_k(\mathbb{K})$ -space. The space V can be recovered as the associated space

$$V \cong \text{Fr}(V) \times_{\text{GL}_k(\mathbb{K})} \mathbb{K}^k. \quad (14.2.23)$$

In this way one can “lift” any natural (that is, $\text{GL}_k(\mathbb{K})$ -equivariant) constructions on the standard spaces \mathbb{K}^k to abstract vector spaces V . For example, $\text{Fr}(V) \times_{\text{GL}_k(\mathbb{K})} \mathbb{K}P^{k-1}$ is the projectivization of V , $\text{Fr}(V) \times_{\text{GL}_k(\mathbb{K})} \wedge^p (\mathbb{K}^k)$ is the exterior algebra of V , etc.

2. The k -fold covering $\mathbb{S}^1 \rightarrow \mathbb{S}^1$ is a principal \mathbb{Z}_k -bundle.
3. If K is a Lie group on which G acts by automorphisms, then $P \times_G K$ is a group bundle. Taking $K = G$ with $G \xrightarrow{\text{Ad}} G$ the action by conjugation, one obtains the associated bundle

$$\text{Ad}_P := P^{[\text{Ad}]} = P \times_G G \quad (14.2.24)$$

called the *adjoint group bundle* of P .

4. If P is a principal G -bundle and $\mathfrak{g} = \text{Lie } G$ with $G \xrightarrow{\text{Ad}} \mathfrak{g}$ being the adjoint representation, then the associated bundle

$$\text{Ad}_P := P^{[\text{Ad}]} = P \times_G \mathfrak{g} \quad (14.2.25)$$

is called the *adjoint vector bundle* of P .

5. Each tensor bundle of a **VB** E , as well as its symmetric and exterior powers can be alternatively defined as the **VBS** associated to the frame bundle $\text{Fr}(E)$ via the corresponding tensor representations of $\text{GL}_k(\mathbb{K})$. For example,

$$\bigwedge^p E \cong \text{Fr}(E)^{[\wedge^p \sigma]} = \text{Fr}(E) \times_{\text{GL}_k(\mathbb{K})} \bigwedge^p \mathbb{K}^k, \quad (14.2.26)$$

where σ is the defining representation of $\text{GL}_k(\mathbb{K})$ on \mathbb{K}^k .

6. Using the same idea for a real **VB** E and $s \in \mathbb{R}$, the line bundles of *scalar s-densities* can be defined as

$$|\det E|^s := \text{Fr}(E)^{[|\det \sigma|^{-s}]} = \text{Fr}(E) \times_{\text{GL}_k(\mathbb{R})} \mathbb{R}, \quad (14.2.27)$$

where the action $|\det \sigma|^{-s}$ multiplies an element of \mathbb{R} by the $(-s)$ -th power of the determinant of the element of $\text{GL}_k(\mathbb{R})$. In particular, sections of $|\det E|^{-1/2}$ are called *half-densities*.

7. The Hopf bundle, which we defined by the free U_1 -action on \mathbb{S}^3 , is therefore a principal U_1 -bundle. Its associated complex line (or real plane) bundle is the tautological line bundle $\gamma_1(\mathbb{C}^2)$ over $\mathbb{CP}^1 \cong \mathbb{S}^2$ (see Exercise 8.2).
8. The “antipodal quotient” of the total space \mathbb{S}^3 of the Hopf bundle (Chern number -1) produces the bundle $SO_3 \rightarrow \mathbb{S}^2$ (Chern number -2), which can be written as $\pi(A) = A_1$, where $A \in SO_3$ and $A_1 \in \mathbb{S}^2 \subset \mathbb{R}^3$ is the first row of A viewed as a unit vector. This bundle is the orthonormal frame bundle of the sphere, $Fr_{SO_3}(T\mathbb{S}^2)$.
9. Recall the principal Stiefel O_k -bundle $V_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n)$ (see (13.8.8)). Its projection map takes an orthonormal k -frame and produces the k -dimensional subspace W of \mathbb{K}^n spanned by that k -frame. Meanwhile this very subspace is the fiber above $W \in G_k(\mathbb{K}^n)$ in the tautological bundle $\gamma_k(\mathbb{K}^n) \rightarrow G_k(\mathbb{K}^n)$. It is obvious that the transition functions of these two bundles coincide. Therefore the tautological bundle is associated to the Stiefel principal bundle:

$$\gamma_k(\mathbb{K}^n) = V_k(\mathbb{K}^n) \times_{O_k} \mathbb{K}^k. \quad (14.2.28)$$

10. The structure group of the tangent bundle $T\mathbb{S}^n$ can be reduced to O_n , and the corresponding orthonormal frame bundle $Fr_{O_n}(T\mathbb{S}^n)$ is the space of pairs (x, \underline{y}) where $x \in \mathbb{S}^n \subset \mathbb{R}^{n+1}$ and \underline{y} is an orthonormal n -frame in \mathbb{R}^{n+1} perpendicular to x . But this space is nothing but $V_{n+1}(\mathbb{R}^{n+1}) = O_{n+1}$. This once again confirms that \mathbb{S}^n is the homogeneous space O_{n+1}/O_n .
11. The Grassmannian of a rank- n **VB** E can be defined as

$$G_k(E) := Fr(E) \times_{GL_n(\mathbb{K})} G_k(\mathbb{K}^n), \quad (14.2.29)$$

where $GL_n(\mathbb{K})$ acts on $G_k(\mathbb{K}^n)$ in the standard way. The fibers of this bundle are the Grassmannians of the corresponding fibers of E , and its sections are nothing but regular distributions on E of rank k .

Similarly, the Stiefel principal bundle of k -frames of E is

$$V_k(E) := Fr(E) \times_{GL_n(\mathbb{K})} V_k(\mathbb{K}^n). \quad (14.2.30)$$

Example 14.2.13 (k -th order frame bundle). Recall Definition 8.2.12 of jet bundles. We can naturally identify the set $J_0^1(\mathbb{R}^m, \mathbb{R}^m)_0^\times$ of invertible one-jets on \mathbb{R}^m with source and target zero with the group $GL_m(\mathbb{R})$. This leads to an interpretation of the linear frame bundle $Fr M = Fr(TM)$ of an m -dimensional manifold M as $J_0^1(\mathbb{R}^m, M)^\times$ of invertible one-jets from \mathbb{R}^m to M with source 0.

More generally, the k -th order frame bundle $Fr^k M$ is defined as $J_0^k(\mathbb{R}^m, M)^\times$, so it consists of k -jets of local charts. Its structure group, as before, is the differential group G_m^k , and the principal right action is given by jet composition from the right. The k -th order tangent

bundle $J_0^k(\mathbb{R}, M)$ is associated to $\text{Fr}^k M$ with respect to the obvious left action of G_m^k on $F = J_0^k(\mathbb{R}, \mathbb{R}^m)$.

Example 14.2.14 (Tangent bundle of a Lie group, part 1). Let G be a real Lie group and let $\pi : P = G \times G \rightarrow G$, $\pi(g_1, g_2) = g_1 g_2^{-1}$, be the principal G -bundle with the “diagonal” principal right G -action: $\Phi_g(g_1, g_2) = (g_1 g, g_2 g)$. It is trivial because there is a global section given by $s : [(g_1, g_2)] \mapsto (g_1 g_2^{-1}, e)$. The Lie subalgebra corresponding to the G -action on P is

$$\mathfrak{h} = \{(A, A) \in \mathfrak{g} \times \mathfrak{g} \mid A \in \mathfrak{g}\}. \quad (14.2.31)$$

Now consider the adjoint representation $G \xrightarrow{\text{Ad}} \mathfrak{g}$ and the associated bundle $P^{[\text{Ad}]} = P \times_G \mathfrak{g}$. Since P is trivial, this bundle is also trivial, and hence isomorphic to the tangent bundle TG . The isomorphism is not unique, but by identifying TG with $G \times \mathfrak{g}$ using the left-invariant Maurer-Cartan form, every such isomorphism can be seen as a global trivialization, hence uniquely determined by a global section of P . The isomorphism corresponding to the global section s defined above is the following:

$$\vartheta : TG \rightarrow P \times_G \mathfrak{g}, \quad (g, X) \mapsto [(g, e), L_{g*}^{-1} X], \quad (14.2.32)$$

$$\vartheta^{-1} : P \times_G \mathfrak{g} \rightarrow TG, \quad [(g_1, g_2), A] \mapsto L_{g_1*} R_{g_2*}^{-1} A. \quad (14.2.33)$$

This example is continued in Example 14.2.19.

Definition 14.2.15 (Natural bundle). A natural bundle F on the category \mathcal{M}_n of n -dimensional smooth manifolds and local diffeomorphisms is a functor assigning to any n -manifold N a **FB** $\pi_N : F(N) \rightarrow N$ and to any local diffeomorphism $f : N_1 \rightarrow N_2$ a bundle morphism $F(f) : F(N_1) \rightarrow F(N_2)$ with base map f , i.e. such that $\pi_{N_2} \circ F(f) = f \circ \pi_{N_1}$.

One often requires the following extra properties:

- (a) F is *local*, i.e. for any inclusion $i : U \hookrightarrow N$ of an open subset, $F(i)$ is the inclusion $p_N^{-1}(U) \hookrightarrow F(N)$.
- (b) F is *regular*, i.e. if M is any smooth manifold and $f : M \times N_1 \rightarrow N_2$ is smooth such that for each $m \in M$ the map $f_m : N_1 \rightarrow N_2$ defined by $f_m(n) := f(m, n)$ is a local diffeomorphism, then the map $M \times F(N_1) \rightarrow F(N_2)$ defined by $(m, p) \mapsto F(f_m)(p)$ is smooth as well. Thus, regularity means that smoothly parametrized families of local diffeomorphisms are transformed into smoothly parametrized families.

It turns out that regularity follows from locality and functoriality.

For example, all tensor bundle functors \mathbb{T}_s^r are natural bundles. The following intuitive characterization of natural bundles holds.

Theorem 14.2.16 ([42, Thm. 1.2.8]). Any local natural bundle on n -dimensional manifolds can be obtained as an associated bundle to some $\text{Fr}^k M$ with respect to a left action of the differential group $G_n^k \curvearrowright F$ on a finite-dimensional manifold F .

The lowest possible choice for k in this theorem is called the *order of the natural bundle*. Notice that the composition of the jet prolongation functor j^k with a r -th order natural bundle is the $(k+r)$ -th order natural bundle $J^k F$.

We now discuss the relationship between equivariant maps on principal bundles and sections of associated bundles. Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a (right) principal G -bundle and (F, G, σ) a (left) G -manifold. Let $\text{Hom}_G(P, F)$ denote the set of smooth equivariant maps $F : P \rightarrow F$, i.e.

$$F \circ \Phi_g = \sigma_{g^{-1}} \circ F. \quad (14.2.34)$$

This definition can be adapted accordingly to the case of a right action σ by replacing $\sigma_{g^{-1}}$ with σ_g . The following proposition shows that there is a natural bijection between such equivariant functions on P and sections of the associated bundle.

Proposition 14.2.17 (Cartan's moving frames [17, Prop. 1.2.6]). For every $\tilde{s} \in \text{Hom}_G(P, F)$ there exists a unique section $s \in \Gamma^\infty(P \times_G F)$ such that the following diagram commutes:

$$\begin{array}{ccc} P & \xrightarrow{\text{id}_P \times \tilde{s}} & P \times F \\ \pi \downarrow & & \downarrow \iota \\ M & \dashrightarrow_{\tilde{s}} & P \times_G F. \end{array} \quad (14.2.35)$$

The assignment $\tilde{s} \mapsto s$ is a bijection from $\text{Hom}_G(P, F)$ onto $\Gamma^\infty(P^{[\sigma]}) = \Gamma^\infty(P \times_G F)$.

Proof. For a given equivariant map \tilde{s} , define

$$s(m) := [(p, \tilde{s}(p))] = \iota_p(\tilde{s}(p)), \quad (14.2.36)$$

where $p \in \pi^{-1}(m)$. This is a well-defined section of $P \times_G F$ because the equivariance property (14.2.34) implies

$$[(\Phi_g(p), \tilde{s}(\Phi_g(p)))] = [(\Phi_g(p), \sigma_{g^{-1}}(\tilde{s}(p)))] = [(p, \tilde{s}(p))] \quad (14.2.37)$$

for all $g \in G$. By definition of s , the diagram commutes. Conversely, since ι_p is a bijection, \tilde{s} can be uniquely reconstructed from s . ■

Remark 14.2.18. Since F is the typical fiber and elements of P are frames for the corresponding fiber of E (cf. Example 13.1.7(4)), \tilde{s} can be thought of as a “moving frame” describing

the section s . For example, if $F = \mathbb{R}^n$, then the values of $\tilde{s}(p)$ on P_m are literally the coordinates of $s(m) \in E_m$ in every possible frame for E_m . The equivariance of \tilde{s} describes the transformations of these coordinates under changes of frames.

More generally, a Cartan moving frame (*repère mobile*) on a G -manifold X is any equivariant function $\varphi : X \rightarrow G$, where G acts on itself by translations. Moving frames exist locally iff the action is free and *regular*. The latter means that the orbits form a regular foliation: every point has a system of arbitrarily small neighborhoods whose intersection with every orbit is connected. Obviously principal bundles are a special case of this. Cartan geometries will be a natural generalization.

Example 14.2.19 (Tangent bundle of a Lie group, part 2). Continuing Example 14.2.14, we expect a correspondence between equivariant \mathfrak{g} -valued functions on $P = G \times G$ and sections of $P \times_G \mathfrak{g}$, which, via the chosen isomorphism ϑ , can be identified with vector fields on G . Let $A \in \mathfrak{g}$ and consider the function

$$\tilde{s}_A : G \times G \rightarrow \mathfrak{g}, \quad (g_1, g_2) \mapsto \text{Ad}_{g_2}^{-1} A. \quad (14.2.38)$$

Clearly this function intertwines the right action $\Phi_g(g_1, g_2) = (g_1g, g_2g)$ with the inverse of the adjoint representation on \mathfrak{g} . For any $g \in G$ we can pick $p = (g, e)$ as the point in the fiber P_g , so the corresponding section of $P^{[\text{Ad}]}$ is

$$s_A(g) = \iota_{(g, e)}(\tilde{s}(g, e)) = [(g, e), A]. \quad (14.2.39)$$

Under the isomorphism $\vartheta : TG \rightarrow P^{[\text{Ad}]}$, this section is identified with the left-invariant vector field generated by A :

$$\vartheta^{-1}(s_A(g)) = L_{g*A} = A_L(g). \quad (14.2.40)$$

Note that the symmetry between left and right was broken when we chose the global section $s(g) = (g, e)$ of P and used it to define the isomorphism ϑ . Had we chosen $s(g) = (e, g)$ instead, the same function \tilde{s}_A would have corresponded to the right-invariant vector field $A_R(g) = \text{Ad}_g^{-1} A_L(g)$. We will continue this work in Example 16.2.6 and use this construction in Example 16.3.21.

14.3 Principal bundle morphisms

We have previously defined morphisms of principal bundles as bundle morphisms modeled on homomorphisms of Lie groups. In the restricted category Prin^G , these morphisms are modeled on id_G and therefore are exactly the G -equivariant maps. Indeed, orbits are exactly the fibers of a principal bundle, and every G -equivariant map $\varphi : P \rightarrow P'$ of right principal G -bundles must map orbits to orbits and be bijective when restricted to one orbit. This proves

that $\text{Hom}_G(P, P') = \text{Mor}_{\text{Prin}^G}(P, P')$.

Now, Proposition 14.2.17 applies, in particular, to the case where $F = P'$ is another right principal G -bundle, and thus yields a bijection between morphisms $\text{Hom}_G(P, P')$ and sections of the associated bundle $P \times_G P' \rightarrow M$. This can be further refined to be seen as a bundle over $M \times M'$. The direct product map $\pi \times \pi' : P \times P' \rightarrow M \times M'$ defined by the bundle projections $\pi : P \rightarrow M$ and $\pi' : P' \rightarrow M'$ descends to a surjective submersion

$$\pi \times_G \pi' : P \times_G P' \rightarrow M \times M'. \quad (14.3.1)$$

This is a **FB** with typical fiber G : given local trivializations (U_α, χ_α) of P and $(U'_\alpha, \chi'_\alpha)$ of P' , one can check that the mapping $\chi_\alpha \times_G \chi'_\alpha$ defined by the commutativity of the square

$$\begin{array}{ccc} \pi^{-1}(U_\alpha) \times \pi'^{-1}(U'_\alpha) & \xrightarrow{(\chi_\alpha, \chi'_\alpha)} & U_\alpha \times U'_\alpha \times (G \times G) \\ \downarrow & & \downarrow \text{id}_{U_\alpha} \times \text{id}_{U'_\alpha} \times \mu \\ (\pi \times_G \pi')^{-1}(U_\alpha \times U'_\alpha) & \dashrightarrow_{\chi_\alpha \times_G \chi'_\alpha} & U_\alpha \times U'_\alpha \times G, \end{array} \quad (14.3.2)$$

with $\mu(g, h) = gh^{-1}$, is a diffeomorphism. Indeed, elements of $P \times_G P'$ should be invariant under the simultaneous right G -action on P and P' and therefore can be thought of as “ratios” $p'p^{-1}$ where $p \in P$ and $p' \in P'$.³³ Clearly such formal combinations are invariant under the right G -action and thus accurately represent elements of $P \times_G P'$. Moreover, from this representation it is obvious that the transition functions in this bundle are exactly the pairs $t_{\beta\alpha} \times t'_{ba} : U_{\alpha\beta} \times U'_{ab} \rightarrow G \times G$ acting on the typical fiber G via $L \times R^{-1}$, so that the transformation law for the local representatives $\{\hat{g}_{\alpha\alpha}\}$ of a (local) section of this bundle reads

$$\hat{g}_{\beta b}(m, m') = t_{\beta\alpha}(m)\hat{g}_{\alpha a}(m, m')t'_{ba}(m')^{-1}. \quad (14.3.3)$$

Now let us further specialize to the case $M = M'$. We can then restrict this bundle over $M \times M$ to the diagonal by defining

$$P \times_{G,M} P' := \Delta^*(P \times_G P'), \quad (14.3.4)$$

where $\Delta : M \hookrightarrow M \times M$ is the diagonal embedding map $m \mapsto (m, m)$. Then $P \times_{G,M} P'$ is an embedded submanifold of $P \times_G P'$ and the induced projection $P \times_{G,M} P' \rightarrow M$ coincides with the restriction of the associated bundle projection $P \times_G P' \rightarrow M$ to this submanifold. Thus $P \times_{G,M} P'$ is an embedded vertical subbundle of the associated bundle $P \times_G P'$. The sections of this new bundle over M have the transformation law

$$\hat{g}_{\beta b}(m) = t_{\beta\alpha}(m)\hat{g}_{\alpha a}(m)t'_{ba}(m)^{-1}. \quad (14.3.5)$$

- Remark 14.3.1.**
1. Note that these bundles aren't principal: the structure group of both $P \times_G P'$ and $P \times_{G,M} P'$ is $G \times G$ despite the fact that the typical fiber is G .
 2. The order in which to quotient by the action of G and restrict to the diagonal doesn't matter, since $P \times_{G,M} P' = \Delta^*((P \times P)/G) = (\Delta^*(P \times P'))/G = P \times_{M,G} P'$. This is even more obvious from comparing the transition functions.

We are finally ready to wrap up the discussion from the beginning of Section 14.2. Recall that we started with a morphism of principal bundles $\vartheta : P \rightarrow P'$ with differing structure groups G, G' covering $\theta : M \rightarrow M'$ and reduced it to a collection of local G' -valued functions $\{\hat{g}'_{\alpha\alpha}\}$ that must satisfy the transformation rule (14.2.2). By now it should be clear that these functions represent a global section of a bundle whose typical fiber is G' and whose structure group is $G \times G'$ acting on G via $((h, h'), g') \mapsto h'g'\lambda(h)^{-1}$. Its transition functions are pairs $m \mapsto (\lambda \circ t_{\beta\alpha}(m), t'_{\beta\alpha}(\theta(m))) \in G \times G'$. By comparing with the definitions of associated bundles and pullbacks we conclude that the functions $\{\hat{g}'_{\alpha\alpha}\}$ define a global section of the bundle

$$P^{[\lambda, \theta]} := P^{[\lambda]} \times_{G', M} (\theta^* P') \rightarrow M. \quad (14.3.6)$$

This bundle can also be written as an associated bundle $(P \times_M \theta^* P') \times_{G \times G'} G'$, or as $(\text{id}_M \times \theta)^*(P \times P') \times_{G \times G'} G'$. All of these are easily seen to be isomorphic by comparing the transition functions. The elements of this bundle can be seen informally as ratios $p'\lambda(p)^{-1}$, where $\pi'(p') = \theta(\pi(p))$, which makes the transformation law (14.2.2) manifest. We thus have

Proposition 14.3.2. *Let $P(G, M)$ and $P'(G', M')$ be PFBs, $\lambda : G \rightarrow G'$ a Lie group homomorphism, and $\theta : M \rightarrow M'$ a smooth map. Then there is a natural bijection between morphisms $P \rightarrow P'$ modeled on λ and covering θ and the set of global sections of the bundle $P^{[\lambda, \theta]}$ and also $G \times G'$ -equivariant G' -valued maps on $P \times_M (\theta^* P')$, where $G \times G'$ acts on G' by $((h, h'), g') \mapsto h'g'\lambda(h)^{-1}$:*

$$\text{Hom}_{\lambda, \theta}(P, P') \cong \Gamma^\infty(P^{[\lambda, \theta]}) \cong \text{Hom}_{G \times G'}(P \times_M \theta^* P, G'). \quad (14.3.7)$$

Example 14.3.3. The existence of a global section is a pretty strong requirement for a general fiber bundle. Proposition 14.3.2 thus shows that there are pretty few morphisms between principal bundles. Let us consider a few common cases.

1. If P' is a bundle over a single point, then it can be identified with a single fiber G' (viewed as a principal homogeneous space), so $\theta^* P' \cong M \times G'$, and we get $P^{[\lambda, \theta]} \cong P^{[\lambda]} \times_{G'} G' \cong P^{[\lambda]}$, which reproduces Proposition 14.2.17 with $F = G'$ acted upon via λ .
2. If $G = G'$ and $\lambda = \text{id}_G$, then $P^{[\lambda, \theta]} = P \times_{G,M} (\theta^* P')$.
3. If $M = M'$ and $\theta = \text{id}_M$, then $P^{[\lambda, \theta]} = P^{[\lambda]} \times_{G', M} P'$.

4. If both $\lambda = \text{id}_G$ and $\theta = \text{id}_M$, then $P^{[\lambda, \theta]} = P \times_{G, M} P'$.

Corollary 14.3.4 ([17, Cor. 1.2.7]). *If $P \xrightarrow{\pi} M$, $P' \xrightarrow{\pi'} M$ are two principal G -bundles over M , then there is a bijection between vertical morphisms from P to P' and sections of $P \times_{G, M} P'$. Namely, if $\vartheta : P \rightarrow P'$ is a G -morphism, the corresponding section is $s_\vartheta(m) = [(p, \vartheta(p))]$ where $p \in \pi^{-1}(m)$.*

This result is very intuitive: G -morphisms $P \rightarrow P'$ are identified with sections of the “bundle of ratios $p'p^{-1}$ ”. Locally, such a section is represented by G -valued functions that prescribe how each fiber should be “twisted”. This also confirms that every G -morphism of bundles over the same base is an isomorphism.

Remark 14.3.5. Proposition 14.2.3 shows that a morphism of principal bundles, together with an equivariant maps of typical fibers, induces a morphism of associated bundles. It is natural to ask whether every bundle morphism can be obtained this way from some principal bundle morphism. Clearly the answer is no because general bundle morphisms need not be modeled on equivariant maps.

Recall that principal bundles are just LieGr-bundles as defined in Definition 7.4.1. One can similarly approach a classification of all morphisms between \mathcal{S} -bundles for arbitrary categories \mathcal{S} of typical fibers. We notice that the maps $\Phi_{\alpha a}$ with transformation law (7.4.8) define a $C^\infty(F, F')$ -valued $(G \times G')$ -equivariant map $\varphi : P \times_M h^*P' \rightarrow C^\infty(F, F')$, where P and P' are the principal bundles associated to E and E' , respectively. The values of this map lie in a $(G \times G')$ -orbit of an element of $\mathcal{S}(F, F')$ (which in the case of principal bundles is the set of homomorphisms $\text{Hom}(G, G')$). Therefore general bundle morphisms can be identified with sections of $(P \times_M h^*P') \times_{G \times G'} \mathcal{S}(F, F')$ (assuming $\mathcal{S}(F, F')$ carries a sufficiently nice topology). This set is usually very different from the set of morphisms between P and P' .

Definition 14.3.6 (Gauge transformations). Let $P \xrightarrow{\pi} M$ be a principal G -bundle. Its group of automorphisms in the category Prin_M^G , i.e. vertical bundle automorphisms, is denoted

$$\text{Gau}(P) := \text{Aut}_M(P), \quad (14.3.8)$$

and is called its group of *active local gauge transformations*, or simply the *gauge group*. By Proposition 14.2.17 and Corollary 14.3.4, there are natural bijections

$$\text{Gau}(P) \cong \Gamma^\infty(P \times_{G, M} P) \cong \text{Hom}_G(P, G) \cong \Gamma^\infty(\text{Ad}(P)), \quad (14.3.9)$$

where $\text{Hom}_G(P, G)$ is the group (via pointwise multiplication) of equivariant G -valued functions on P intertwining the right G -action on P with the action of G on itself by conjugations Ad (i.e. $\sigma = \text{Ad}$ in (14.2.34)).

Note that in the general category Prin^∞ , the automorphism group $\text{Aut}(P)$ also contains automorphisms that cover nontrivial diffeomorphisms of the base. Thus we have the exact sequence

$$1 \rightarrow \text{Gau}(P) \rightarrow \text{Aut}(P) \rightarrow \text{Diff}(M). \quad (14.3.10)$$

The last map is generally not surjective.

Example 14.3.7 (Deck transformations). If $P \xrightarrow{\pi} M$ is a regular covering map, then its group of deck transformations $G = \text{Aut}(\pi)$ is exactly the gauge group of P as a principal G -bundle.

Remark 14.3.8.

- It might seem strange that we have identified a group with two spaces of sections, because most bundles don't have any canonical section that can be identified with the identity element of a group. However, since the structure group of $P^{[\text{Ad}]}$ acts by conjugation, it preserves the identity of the typical fiber G , and thus this bundle does have a natural "identity section". Similarly, $P \times_{G,M} P$ has the same kind of section given by $m \mapsto [(p, p)]$ where $p \in \pi^{-1}(m)$.
- In physics, the phrase "gauge transformation" usually refers to the local gauge transformations, which are just different choices of a G -atlas. By selecting a local frame (section) $s_U : U \rightarrow P$ we can identify every global gauge transformation $\varphi \in \text{Hom}_G(P, G)$ with a local G -valued function $\varphi \circ s_U : U \rightarrow G$. However, unlike these local functions, the equivariant map φ is an intrinsic quantity of the bundle.

14.4 Bundle reduction

First we recall our definition of reductions of principal bundles.

Definition 14.4.1 (Bundle reduction). A vertical morphism (ϑ, λ) of principal bundles $Q(M, H)$ and $P(M, G)$ is called a λ -reduction, or simply a reduction of P to H , if Q is a subbundle of P . That is, if ϑ and λ are injective immersions. In this case, P is called λ -reducible to H and Q is called a λ -reduction of P .

First we characterize bundle reductions in terms of transition functions.

Proposition 14.4.2 ([17, Prop. 1.6.1]). *If $\lambda : H \rightarrow G$ is an injective Lie group homomorphism, then a principal bundle $P(M, G)$ is λ -reducible iff it admits a cocycle taking values in $\text{im}(\lambda)$.*

Proof. Let $Q(M, H)$ be a λ -reduction of P and let (ϑ, λ) be the corresponding injective morphism. Let $\{(U_\alpha, \chi_\alpha)\}$ be a bundle atlas of Q with transition functions $\{t_{\beta\alpha}\}$. Since ϑ maps local sections of Q to local sections of P , each χ_α defines a local trivialization χ'_α of P over U_α . Let κ_α and κ'_α be the equivariant maps corresponding to χ_α and χ'_α , respectively. One can check that $\lambda \circ \kappa_\alpha = \kappa'_\alpha \circ \vartheta$. Hence, the transition maps of $\{\chi'_\alpha\}$ are

$$t'_{\beta\alpha}(m) = \kappa'_\beta(\vartheta(q))\kappa'_\alpha(\vartheta(q))^{-1} = \lambda(\kappa_\beta(q))\lambda(\kappa_\alpha(q))^{-1} = \lambda(t_{\beta\alpha}(m)). \quad (14.4.1)$$

Conversely, if $\{t'_{\beta\alpha}\}$ is a cocycle of P that takes values in $\text{im}(\lambda)$, it defines H -valued functions $\{t_{\beta\alpha}\}$ via $t'_{\beta\alpha} = \lambda \circ t_{\beta\alpha}$. Since injective Lie group homomorphisms are immersions, (H, λ) is a Lie subgroup. Since Lie subgroups are weakly embedded, $t_{\beta\alpha}$ are smooth. They satisfy the cocycle property and thus determine a principal H -bundle Q . The morphism ϑ can be defined via its local representatives $\vartheta_\alpha : \pi_Q^{-1}(U_\alpha) \rightarrow \pi_P^{-1}(U_\alpha)$:

$$\vartheta_\alpha := \chi_\alpha'^{-1} \circ (\text{id}_{U_\alpha} \times \lambda) \circ \chi_\alpha. \quad (14.4.2)$$

Since $t'_{\beta\alpha} = \lambda \circ t_{\beta\alpha}$, we have $\vartheta_\alpha = \vartheta_\beta$ for any $U_{\alpha\beta} \neq \emptyset$. Thus, the family $\{\vartheta_\alpha\}$ defines an equivariant map $\vartheta : Q \rightarrow P$. By construction, (ϑ, λ) is a λ -reduction. ■

The second criterion for reducibility is in terms of equivariant functions.

Proposition 14.4.3 ([17, Prop. 1.6.2]). *Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a principal bundle and $G \overset{\sigma}{\curvearrowright} F$ a transitive Lie group action. Let $f \in F$ and let $G_f < G$ be the stabilizer group of f under the action σ . Then every equivariant map $\varphi \in \text{Hom}_G(P, F)$ defines a reduction of P to an embedded principal G_f -subbundle*

$$Q_f := \varphi^{-1}(f) = \{p \in P : \varphi(p) = f\}. \quad (14.4.3)$$

Conversely, every such reduction defines an element $\varphi \in \text{Hom}_G(P, F)$.

Proof. Since σ is transitive and φ is equivariant, φ is a submersion. Hence, by the Level Set Theorem 3.9.7, $Q_f = \varphi^{-1}(f)$ is an embedded submanifold of P for all $f \in F$. Since for every $q \in Q_f$ and every $g \in G_f$,

$$\varphi(\Phi_g(q)) = \sigma_{g^{-1}}(\varphi(q)) = \sigma_{g^{-1}}(f) = f, \quad (14.4.4)$$

Q_f is G_f -invariant. Thus Φ induces a free right G_f -action on Q_f , denoted by the same symbol. Since G_f is closed (Proposition 13.3.3), the induced action is proper (Proposition 13.5.4). Hence Q_f is a principal G_f -bundle over Q_f/G_f . The natural inclusions $Q_f \hookrightarrow P$ and $G_f \hookrightarrow G$ define a principal bundle morphism. Let $\theta : Q_f/G_f \rightarrow M$ be the corresponding projection. It remains to show that θ is a diffeomorphism. By local triviality, it suffices to show bijectivity. For that purpose, we show that Q_f intersects every fiber of P exactly along the G_f -orbits. Let $m \in M$ and $p \in \pi^{-1}(m)$. Since σ is transitive, there exists a $g \in G$ such that $\varphi(p) = \sigma_g(f)$. Then $\varphi(\Phi_g(p)) = \sigma_{g^{-1}}(\varphi(p)) = f$, that is, $\Phi_g(p) \in Q_f$, so Q_f intersects with the fiber over m . To compute this intersection, let $q_1, q_2 \in Q_f$ and $g \in G$ such that $q_2 = \Phi_g(q_1)$. Then

$$f = \varphi(q_2) = \varphi(\Phi_g(q_1)) = \sigma_{g^{-1}}(\varphi(q_1)) = \sigma_{g^{-1}}(f), \quad (14.4.5)$$

that is, $g \in G_f$. Thus the fiber of Q_f is exactly the G_f -orbit, and Q_f is a reduction of P to the subgroup G_f .

Conversely, let $Q < P$ be a reduction of P with structure group $G_f < G$ and with the morphism given by the natural inclusion. Then we take the constant map $\varphi : Q \rightarrow F$, $\varphi(q) := f$, and extend it to a map $\varphi : P \rightarrow F$ by equivariance:

$$\varphi(\Phi_g(q)) = \sigma_{g^{-1}}(f), \quad g \in G, q \in Q. \quad (14.4.6)$$

This map is well-defined: if $\Phi_{g_1}(q_1) = \Phi_{g_2}(q_2)$ for $q_1, q_2 \in Q$, then $g_1 g_2^{-1} \in G_f$ and thus

$$\sigma_{g_2^{-1}}(f) = \sigma_{g_1^{-1}} \circ \sigma_{g_1 g_2^{-1}}(f) = \sigma_{g_1^{-1}}(f). \quad (14.4.7)$$

By construction, φ is smooth and equivariant. ■

The element f should be thought of as the “model G -invariant structure”, and φ as encoding the particular way of copying that model structure to every fiber of the bundle. This will make more sense in particular examples.

Remark 14.4.4 ([17, Rem. 1.6.3]). The reduction Q_f depends on the choice of $f \in F$ as follows: for every $f' \in F$ there is a $g \in G$ such that $f' = \sigma_g(f)$. Then for every $q \in Q_f$ we have

$$\varphi(\Phi_{g^{-1}}(q)) = \sigma_g(\varphi(q)) = \sigma_g(f) = f'. \quad (14.4.8)$$

Thus $Q_{f'} = \Phi_{g^{-1}}(Q)$. Moreover, the corresponding structure group is $G_{f'} = G_{\sigma_g(f)} = gG_fg^{-1}$.

The following proposition characterizes the isomorphism classes of bundle reductions to a given structure group G_f .

Proposition 14.4.5 ([17, Prop. 1.6.4]). *Let $\varphi \in \text{Hom}_G(P, F)$ and let $\vartheta \in \text{Gau}(P)$ be a vertical automorphism of P . If φ defines the bundle reduction Q_f , then $\varphi \circ \vartheta$ defines the bundle reduction $\vartheta^{-1}(Q_f)$. In particular, two reductions to the structure group G_f are equivalent iff the defining equivariant maps are related by a vertical automorphism.*

Proof. The reduced bundle defined by $\varphi \circ \vartheta$ is

$$\{p \in P : \varphi \circ \vartheta(p) = f\} = \vartheta^{-1}(\{p \in P : \varphi(p) = f\}) = \vartheta^{-1}(Q_f). \quad (14.4.9)$$

■

Proposition 14.4.3 implies a useful characterization of reductions of a principal bundle $P(M, G)$ to a given closed subgroup H if G . To formulate it, we consider the natural action of G on the homogeneous space G/H by left translation and build the associated bundle $P \times_G G/H$, whose total space is P/H .

Corollary 14.4.6 ([17, Cor. 1.6.5]). *The reductions of a principal G -bundle P to a closed subgroup $H \triangleleft G$ are in one-to-one correspondence with the smooth sections of the associated bundle $P \times_G G/H \cong P/H$.*

Proof. By Proposition 14.2.17, the sections of $P \times_G G/H$ are in one-to-one correspondence with elements of $\text{Hom}_G(P, G/H)$. Since G/H is a transitive G -manifold, we can apply Proposition 14.4.3 with $f = [e]$. ■

Example 14.4.7 ([17, Ex. 1.6.6]). 1. Let E be a \mathbb{K} -vector bundle of rank k . Recall that E is orientable iff it admits a bundle atlas with transition functions of positive determinant. Equivalently, E is orientable iff it admits a nowhere vanishing section of the determinant line bundle $\det E = \wedge^k E^*$. Thus, an orientation of E may be viewed as a section of the associated bundle

$$\text{Fr}(E) \times_{\text{GL}_k(\mathbb{K})} (\text{GL}_k(\mathbb{K})/\text{GL}_k^+(\mathbb{K})). \quad (14.4.10)$$

Now, Corollary 14.4.6 implies that E is orientable iff $\text{Fr}(E)$ is reducible to $\text{GL}_k^+(\mathbb{K})$.

2. Let E be a \mathbb{K} -vector bundle of rank n endowed with a fiber metric h (sesquilinear in the cases $\mathbb{K} = \mathbb{C}, \mathbb{H}$). It need not be positive-definite.

(a) If $\mathbb{K} = \mathbb{R}$, then a bundle metric can be viewed as a section of the associated bundle

$$\text{Fr}(E) \times_{\text{GL}_n(\mathbb{R})} S^2(\mathbb{R}^n)^*, \quad (14.4.11)$$

where S^2 denotes the symmetric tensor square. By the Sylvester Theorem, $\text{GL}_n(\mathbb{R})$ acts transitively on the subspace $S^2_{(k,l)}(\mathbb{R}^n)^* < S^2(\mathbb{R}^n)^*$ consisting of elements of rank $n = k + l$ and signature (k, l) and the stabilizer of the element $\eta = I_k \oplus (-I_l)$ is $\text{O}_{k,l}$. Thus

$$\text{GL}_n(\mathbb{R})/\text{O}_{k,l} \cong S^2_{(k,l)}(\mathbb{R}^n)^*, \quad (14.4.12)$$

and E admits a fiber metric with signature (k, l) iff $\text{Fr}(E)$ is reducible to a principal $\text{O}_{k,l}$ -bundle. The latter is the bundle of orthonormal frames $\text{Fr}_{\text{O}_{k,l}}(E, \mathsf{h})$. Note that specifying such a reduction is equivalent to specifying a bundle metric (section of the associated bundle) according to Proposition 14.4.2, or alternatively a metric (f in the language of Proposition 14.4.3) on the model fiber plus an equivariant $S^2_{(k,l)}(\mathbb{R}^n)^*$ -valued (metric-valued) function on the frame bundle.

(b) If $\mathbb{K} = \mathbb{C}$, then a fiber metric can be seen as a section of the associated bundle

$$\text{Fr}(E) \times_{\text{GL}_n(\mathbb{C})} (\bar{\mathbb{C}}^n)^* \xrightarrow{s} (\mathbb{C}^n)^*, \quad (14.4.13)$$

where \otimes^s is the symmetric tensor product. By the Sylvester Theorem, $\text{GL}_n(\mathbb{C})$ acts transitively on the subset of nondegenerate elements in $(\bar{\mathbb{C}}^n)^* \otimes^s (\mathbb{C}^n)^*$ with stabilizer U_n . Thus E admits a Hermitean fiber metric iff $\text{Fr}(E)$ is reducible to a principal U_n -bundle, which then coincides with the bundle of unitary frames $\text{Fr}_{\text{U}_n}(E)$.

The following proposition shows that principal bundle reductions do not change the isomorphism class of associated bundles.

Proposition 14.4.8 ([17, Prop. 1.6.7]). *Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a PFB and let $G \curvearrowright F$ be a Lie group action. Let Q be a reduction of P to the structure group H defined by the morphism (ϑ, λ) . Let $H \curvearrowright F$ be the associated H -action. Then the associated FBS $P \times_G F$ and $Q \times_H F$ are isomorphic.*

Proof. Proposition 14.4.2 implies that P has a bundle atlas with transition functions taking values in $\text{im}(\lambda)$. Since λ is an immersion, it turns H into a weakly embedded subgroup of G , so P and Q can be taken to have the same H -valued transition functions, and the action $\sigma \circ \lambda$ is simply the restriction of σ to H . Hence their associated bundles also have the same transition functions and are therefore isomorphic. ■

We have previously defined G -structures as equivalence classes of G -atlases on a given FB. Since every general FB (with no structure) is associated to a principal $\text{Diff}(F)$ -bundle, it is now obvious that a G -structure is simply a bundle reduction.

Definition 14.4.9 (G -structure). Given a locally trivial FB $E \xrightarrow{\pi} M$ with typical fiber F , and a faithful action $G \curvearrowright F$ (so that G can be seen as a Lie subgroup of $\text{Diff}(F)$), a G -structure on E is a bundle reduction Q of the associated principal $\text{Diff}(F)$ -bundle $\mathcal{P}(E)$ to G , together with the induced isomorphism $E \cong Q \times_G F$.

As we have shown, every reduction of $\mathcal{P}(E)$ to G can be identified with a section of the bundle $\mathcal{P}(E)/G$, which is equivalent to picking a G -equivalence class of frames over each point of the base.

Now that we have come full circle and finished defining fiber bundles – a task that we started in Section 7 – we are ready to start looking at more sophisticated additional structures on them. These structures, in turn, will help us find quantitative criterions for the existence of

certain G -structures on a given bundle, and eventually classify all bundles in terms of certain topological invariants.

15 Klein Geometry

15.1 Klein geometry

Felix Klein saw geometry as the study of invariants under actions of Lie groups. This led him to start the famous Erlangen program for the classification of all “symmetric” geometries corresponding to simple Lie groups. He noticed that each interesting symmetric geometry is based on a connected manifold with a *transitive* action of a Lie group G of “motions”, and moreover, all properties of figures studied in the geometry remain invariant under these motions. For example, in the case of Euclidean geometry, the properties studies are angle and length, and the group is the group of rigid motions. Thus Klein unified a wide range of classical geometries by defining a geometry as consisting of a Lie group G , a smooth manifold M with a smooth, faithful, and transitive action of G . This means that Klein geometries are a special kind of principal bundles. We shall not require faithfulness in our treatment here.

Let $H \triangleleft G$ be a closed Lie subgroup and consider the action $H \curvearrowright G$ by right translation. Let K be the maximal normal subgroup of G lying inside H . By the Initial Subgroup Theorem 12.7.2, K is a weakly embedded Lie subgroup of G . Since H is closed, the closure \bar{K} is also a normal subgroup of G lying inside H . By maximality, we conclude that K is closed, and hence embedded, so $K \triangleleft H \triangleleft G$. By Corollary 13.7.3(1), the canonical map $\pi : G/K \rightarrow G/H$ is a PFB with fiber H/K , inducing a diffeomorphism

$$(G/K)/(H/K) \rightarrow G/H \tag{15.1.1}$$

commuting with the canonical left G/K -actions.

Definition 15.1.1 (Klein geometry). A Klein geometry is a pair (G, H) , where G is a Lie group and $H \triangleleft G$ is a closed subgroup such that the homogeneous space G/H is connected. G is called the *principal group* of the geometry. The *kernel* of the Klein geometry is the maximal normal subgroup K of G lying inside H . The Klein geometry is called *effective* (or *faithful*) if $K = \{e\}$ and *locally effective* if K is discrete. A Klein geometry is *geometrically oriented* if G is connected. The homogeneous space $M = G/H$ is called the *space of the Klein geometry*. It is called *reductive* if the adjoint representation of G viewed as a $\text{Ad}(H)$ -module decomposes into a direct sum $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, where \mathfrak{g} and \mathfrak{h} are the Lie algebras of G and H , respectively.

Remark 15.1.2. 1. The same space $M = G/H$ might be the space of multiple Klein geometries, so a Klein geometry is not just a homogeneous space, but rather a principal bundle of the form $G \rightarrow G/H$.
2. Every Klein geometry is a *tautological bundle* in the sense that the fiber over $gH \in G/H$ is the coset $gH \subset G$ itself.

If (G, H) is a Klein geometry with kernel K , then, for any closed subgroup $N < K$ that is normal in G , the pair $(G/N, H/N)$ is also a Klein geometry with the same space $(G/N)/(H/N) \cong G/H$. Of course, these geometries are all ineffective except when $N = K$. This leads to the following.

Definition 15.1.3 (Associated effective Klein geometry). if (G, H) is a Klein geometry with kernel K , then $(G/K, H/K)$ is called the associated effective Klein geometry.

More generally, if K_0 is the identity component of K , then $(G/K_0, H/K_0)$ is a locally effective Klein geometry with kernel K/K_0 .

Remark 15.1.4. Despite the space of the Klein geometry being “independent” of its kernel, ineffective Klein geometries are still important to study, seeing as many important phenomena such as spin can only be described as ineffective geometries. Still, all geometrically important examples of Klein geometries are locally effective. Nevertheless, non-locally effective descriptions still often show up because they are easier to formulate.

The next result shows that every Klein geometry (G, H) determines a geometrically oriented geometry (G_0, H') with the same space but a possibly smaller principal group.

Proposition 15.1.5 ([41, Prop. 3.5]). *Let (G, H) be a Klein geometry. Let G_0 be the identity component of G and set $H' := H \cap G_0$. Then*

- (i) $G = G_0 \cdot H$,
- (ii) $G/H = G_0/H'$.

Proof. (i) It suffices to show that $G \subset G_0 \cdot H$. Let $g \in G$. Since G/H is connected, there is a path $gH \sim eH$. Since the map $G \rightarrow G/H$ is a bundle, we can lift this path to a path $g \sim h$ where $h \in H$. Thus $gh^{-1} \in G_0$ and $g = (gh^{-1})h \in G_0 \cdot H$.

(ii) By (i) the map $j : G_0 \rightarrow G/H$ is surjective. Suppose $g_1, g_2 \in G_0$. Then $g_2^{-1}g_1 \in G_0$ and

$$j(g_1) = j(g_2) \Leftrightarrow g_1 \in g_2H \Leftrightarrow g_2^{-1}g_1 \in H. \quad (15.1.2)$$

Thus j induces a diffeomorphism $G_0/H' \rightarrow G/H$. ■

Example 15.1.6. 1. $(\text{Aff}_n(\mathbb{K}), \text{GL}_n(\mathbb{K}))$ is the affine geometry on \mathbb{K}^n . Similarly, $(\text{Aff}_n^+(\mathbb{K}), \text{GL}_n^+(\mathbb{K}))$ is the affine geometry with an orientation.
 2. $(E_n^+(\mathbb{R}), \text{SO}_n)$ is the “school geometry” of orientation-preserving congruences, or rigid motions, on \mathbb{R}^n . This adds a metric to the real affine geometry.
 3. Let H be the subgroup of $\text{GL}_n^+(\mathbb{R})$ consisting of matrices A such that $AA^T = \lambda I$ for some $\lambda > 0$. Every such matrix has the form $\sqrt{\lambda}R$ where $R \in \text{SO}_n$ is a rotation matrix, so this is precisely the group of similarity transformations in Euclidean geometry. Then the Klein geometry $(H \rtimes \mathbb{R}^n, H)$ can be called the “geometry of similarity” on \mathbb{R}^n .

4. $(\mathrm{SL}_2(\mathbb{R}), \mathrm{SO}_2)$ is the *hyperbolic half-plane* $M = \{z \in \mathbb{C} \mid \mathrm{Im} z > 0\}$ with the transitive action of Möbius transformations $\mathrm{SL}_2(\mathbb{R})$ acting by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}. \quad (15.1.3)$$

5. Let $H < \mathrm{GL}_{n+1}(\mathbb{R})$ be the subgroup

$$H = \left\{ \begin{pmatrix} \lambda & v \\ 0 & A \end{pmatrix} \middle| A \in \mathrm{GL}_n(\mathbb{R}), v \in \mathbb{R}^n, \lambda \in \mathbb{R}^\times \right\}. \quad (15.1.4)$$

Then $(\mathrm{GL}_{n+1}(\mathbb{R}), H)$ is the *projective geometry* on $\mathbb{R}P^n$ with the standard action of $\mathrm{GL}_{n+1}(\mathbb{R})$ on lines in \mathbb{R}^{n+1} .

6. (SO_{n+1}, O_n) , with O_n embedded via

$$A \mapsto \begin{bmatrix} \det A^{-1} & 0 \\ 0 & A \end{bmatrix}, \quad (15.1.5)$$

is the elliptic geometry on $\mathbb{R}P^n$ with Euclidean transformations acting on lines in \mathbb{R}^{n+1} . This adds a metric to the projective geometry.

7. $(O_{n,1}^+ \rtimes \mathbb{R}^n, O_{n,1}^+)$ is the *Minkowski geometry* on \mathbb{R}^n .
8. Recall that under the standard Minkowski metric of signature $(n+1, 1)$ on \mathbb{R}^{n+2} , the light cone $C = \{x \mid \|x\| = 0\}$ is homogeneous in the sense that if $x \in C$, then $\lambda x \in C$ for all $\lambda \in \mathbb{R}$. The set $C/\sim \cong \mathbb{S}^n$ of null lines through the origin thus inherits a transitive $O_{n+1,1}$ -action with stabilizers isomorphic to the Poincaré conformal group $P \cong CO_n \rtimes \mathbb{R}^n$. This is the *conformal geometry* on \mathbb{S}^n . In the case $n = 2$ this is the *Möbius sphere geometry*.
9. If $H < G = O_{2,2}^+$ (where the metric is $\mathrm{diag}(-1, 1, 1, -1)$) is the stabilizer of the line spanned by $(1, 1, 0, 0)$, then $Q = G/H$ can be identified with the *projective completion* of a 1-sheet hyperboloid in \mathbb{R}^3 given by $x_1^2 + x_2^2 - x_3^2 = 1$, and so is diffeomorphic to the torus \mathbb{T}^2 . This is called the *Lie sphere geometry* on the *Lie quadric* Q . In this geometry lines, circles and points are treated on equal footing. This generalizes to $O_{n+2,2}$.
10. $(\mathrm{SL}_2(\mathbb{C}), H)$, where H is the stabilizer of a point of \mathbb{CP}^1 , then this is the geometry of a complex projective line on \mathbb{S}^2 and $\mathrm{SL}_2(\mathbb{C})$ is the group of Möbius transformations.

15.2 Locally Klein geometry

Much like with principal bundles, we shall follow up the original geometric definition by a bundle definition. To motivate it, we recall the following natural data associated to a Klein geometry (G, H) :

- (a) the connected homogeneous manifold $M = G/H$;
- (b) the principal H -bundle $G \rightarrow G/H$;
- (c) the Maurer-Cartan form $\Theta_G : TG \rightarrow \mathfrak{g}$ satisfying

- (i) Θ_G is a linear isomorphism on each fiber,
- (ii) $R_h^* \Theta_G = \text{Ad}_h^{-1} \Theta_G$ for all $h \in H$,
- (iii) $\Theta_G(X_L) = X$ for any $X \in \mathfrak{h}$ and corresponding left-invariant vector field X_L .

These vector fields should be thought of as H -invariant *vertical* (tangent to the fibers) vector fields on the principal bundle.

The notion of isomorphism of Klein geometries is exactly that of isomorphisms of PFBs, with the extra requirement that they be Lie group isomorphisms (which amounts to the preservation of identity).

Definition 15.2.1 (Isomorphic Klein geometries). An isomorphism between two Klein geometries (G, H) and (G', H') is a Lie group homomorphism $\vartheta : G \rightarrow G'$ such that $\vartheta(H) = H'$.

In particular, $(G, H) \cong (G, gHg^{-1})$. Another useful notion is that of *mutation*, which preserves the structure of the Klein geometry only in an infinitesimal sense near the point $[e] \in G/H$. While an isomorphism is a pair of Lie group homomorphisms (ϑ, λ) where $\lambda = \vartheta|_H$, a mutation replaces ϑ with a linear isomorphism of Lie algebras which agrees with λ in a natural way.

Definition 15.2.2 (Mutation of Klein geometry). A mutation of a Klein geometry is a pair of Klein geometries (G, H) and (G', H') and a pair (ϕ, λ) consisting of a Lie group isomorphism $\lambda : H_1 \rightarrow H_2$ and an isomorphism of vector spaces $\phi : \mathfrak{g} \rightarrow \mathfrak{g}'$ which extends the induced isomorphism $\lambda_{*e} : \mathfrak{h} \rightarrow \mathfrak{h}'$ as an $\text{Ad}(H)$ -equivariant map, i.e.

$$\phi(\text{Ad}_h X) = \text{Ad}_{\lambda(h)} \phi(X), \quad \phi|_{\mathfrak{h}} = \lambda_{*e}. \quad (15.2.1)$$

Exercise 15.1. Show that $(E_n^+(\mathbb{R}), \text{SO}_n(\mathbb{R}))$, $(\text{SO}_{n+1}, \text{SO}_n)$, and $(\text{SO}_{n,1}, \text{SO}_n)$ are all mutations of each other (they are Euclidean space \mathbb{R}^n , elliptic space \mathbb{S}^n , and hyperbolic space $H^n \cong \mathbb{R}^n$, respectively).

Definition 15.2.3 (Locally Klein geometry). A locally Klein geometry is a triple (Γ, G, H) , where $H \triangleleft G$ is a closed subgroup and $\Gamma \triangleleft G$ is a discrete subgroup such that its action by left translations on the homogeneous space G/H is faithful, and such that the double coset space $\Gamma \backslash G/H$ is connected. It is called *geometrically oriented* if G is connected.

Note that we do not require G/H to be connected. However the following exercise shows that we can always replace (Γ, G, H) by a locally Klein geometry (Γ', G', H) with the same space, for which (G', H) is a Klein geometry.

Exercise 15.2. Let (Γ, G, H) be a Klein geometry and set $G' = G_0 \cdot H$, where G_0 is the identity component, and $\Gamma' = G' \cap \Gamma$. Show that G' is a Lie group, G'/H is connected, and the inclusion $G' \hookrightarrow G$ induces a diffeomorphism $\Gamma' \backslash G' \rightarrow \Gamma \backslash G$.

Locally Klein geometries can also be defined as principal bundles. This depends on the following lemma.

Lemma 15.2.4 ([41, Lem. 3.12]). *Let Γ and H be Lie groups with free commuting left and right proper actions (respectively) on a smooth manifold P . Then*

$$(\Gamma \backslash P) \times H \rightarrow \Gamma \backslash P \text{ is proper} \Leftrightarrow \Gamma \times (P/H) \rightarrow P/H \text{ is proper.} \quad (15.2.2)$$

Proof. By symmetry it suffices to prove only the forward implication.

Step 1. The simultaneous action $((\Gamma \times H) \times P \rightarrow P$ by $((g, h), p) \mapsto gph^{-1}$ is proper. Let $A, B \subset P$ be compact and let $C = \{(g, h) \in \Gamma \times H \mid gAh \cap B \neq \emptyset\}$. We must show that C is compact, and it suffices to show it is sequentially compact (every sequence has a convergent subsequence).

Let $(g_i, h_i) \in \Gamma \times H$ be a sequence in C . Let A' and B' be the (compact) images of A and B , respectively, in $\Gamma \backslash P$. Since the H -action on $\Gamma \backslash P$ is proper by assumption, the set $C' = \{h \in H \mid A'h \cap B' \neq \emptyset\}$ is compact. Moreover it is clear that $h_i \in C'$, hence (h_i) has a convergent subsequence. Replacing $(g_i, h_i) \in C$ by the corresponding subsequence, we may assume that

$$\lim h_i = h_\infty \in C'. \quad (15.2.3)$$

Since $(g_i, h_i) \in C$ we may choose points $a_i \in A$ and $b_i \in B$ such that $g_i a_i h_i = b_i$. Since A and B are compact, (a_i) and (b_i) have convergent subsequences. Replacing $(g_i, h_i) \in C$ by a corresponding subsequence, we obtain

$$\lim a_i = a_\infty \in A, \quad \lim b_i = b_\infty \in B. \quad (15.2.4)$$

Now let $K \subset P$ be a compact neighborhood of $a_\infty h_\infty$ so that $a_i h_i \in K$ for sufficiently large i . Since $\Gamma \times P \rightarrow P$ is a proper action, the set

$$C'' = \{g \in \Gamma \mid gK \cap \{b_\infty\} \neq \emptyset\} = \{g \in \Gamma \mid g^{-1}b_\infty \in K\} \quad (15.2.5)$$

is compact. Now $\lim g_i^{-1}b_\infty = \lim g_i^{-1}b_i = \lim a_i h_i = a_\infty h_\infty \in K$. Thus $g_i^{-1}b_\infty \in K$ for sufficiently large i . Thus (g_i) has a convergent subsequence, and we may pass to this subsequence to obtain

$$\lim g_i = g_\infty \in \Gamma. \quad (15.2.6)$$

It follows that the original sequence $(g_i, h_i) \in \Gamma \times H$ has a convergent subsequence, and hence C is sequentially compact.

Step 2. The action $\Gamma \times (P/H) \rightarrow P/H$ is proper.

Let $A, B \subset P/H$ be compact. We must show that $\{g \in \Gamma \mid gA \cap B \neq \emptyset\}$ is compact. If we write $A = \bigcup_{1 \leq i \leq a} A_i$, $B = \bigcup_{1 \leq j \leq b} B_j$, with A_i and B_j compact, then it suffices to show that $\{g \in \Gamma \mid gA_i \cap B_j \neq \emptyset\}$ is compact for all i, j . Thus it suffices to assume that A and B are small sets (i.e. subordinate to some open covering of X/H). Since $P \rightarrow P/H$ is a principal H -bundle, there is an open covering of P/H over which P is trivialized. We may assume that A and B are subordinate to this cover.

Applying a local section to A and B , we can obtain compact sets A' and B' in P with images A and B in P/H . Thus, $\{(g, h) \in \Gamma \times H \mid gA'h^{-1} \cap B' \neq \emptyset\}$ is compact. Therefore the image of this set under the canonical projection $\Gamma \times H \rightarrow \Gamma$ is also compact; but this image is clearly $\{g \in \Gamma \mid gA \cap B \neq \emptyset\}$. ■

This means that if P has two commuting free and proper group actions on it as in the above theorem, we have the following commuting square of principal bundles:

$$\begin{array}{ccc} P & \longrightarrow & P/H \\ \downarrow & & \downarrow \\ \Gamma \backslash P & \longrightarrow & \Gamma \backslash P/H, \end{array} \quad (15.2.7)$$

where the left and right arrows are principal Γ -bundles and the top and bottom arrows are principal H -bundles.

Corollary 15.2.5 ([41, Cor. 3.13]). *If (Γ, G, H) is a locally Klein geometry, then the quotient map $\Gamma \backslash G \rightarrow \Gamma \backslash G/H$ is a principal H -bundle.*

Proof. Apply Lemma 15.2.4, taking $P = G$, and use the left and right translation actions of Γ and H , respectively. Since Γ and H are closed subgroups, these actions are proper. Since (Γ, G, H) is locally klein, it follows that the action $\Gamma \times (G/H) \rightarrow G/H$ is free and proper. So by the lemma, $(\Gamma \backslash G) \times H \rightarrow \Gamma \backslash G$ is also free and proper. ■

According to the last exercise, the principal bundle $\Gamma \backslash \Gamma \rightarrow \Gamma \backslash G/H$ is the same as the principal bundle $\Gamma' \backslash G' \rightarrow \Gamma' \backslash G'/H$, where $G' = G_0 \cdot H$ and $\Gamma' = G' \cap \Gamma$.

Since the Maurer-Cartan form Θ_G on G is left-invariant, it induces a form $\Theta_{\Gamma \backslash G}$ on $\Gamma \backslash G$ with all the same properties.

Example 15.2.6. Consider the hyperbolic plan $M = \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2$. It is known that any closed (compact with no boundary), orientable surface M_g of genus $g > 1$ can appear as the quotient $\Gamma \backslash M$ for some choice of $\Gamma < \mathrm{SL}_2(\mathbb{R})$. Let Γ_g be such a subgroup. Then $M_g = \Gamma_g \backslash M$ is a locally Klein geometry.

Example 15.2.7 (Monopole bundles). Consider the Hopf bundle $\mathbb{S}^3 \rightarrow \mathbb{S}^2$ defined as a U_1 -principal bundle, or as a Klein geometry of type (SU_2, U_1) . Let $m \geq 1$ and $\Gamma = \mathbb{Z}_m \subset U_1$ act on \mathbb{S}^3 by the restriction of the U_1 action. The space $\Gamma \backslash \mathbb{S}^3$ is the loop space $L(m; 1)$. We get the principal U_1 -bundle

$$L(m; 1) \rightarrow L(m; 1)/U_1 = \mathbb{S}^2 \quad (15.2.8)$$

(since the induced action of U_1 on $L(m; 1)$ is trivial). Note that the structure groups here are abelian so right and left actions are indistinguishable.

In the standard atlas consisting of two hemispheres (Example 7.7.8), the transition function of this new bundle is $\tau_+(z) = z^{-m}$, i.e. this is exactly the U_1 -bundle with Chern number $-m$ from Example 7.7.9. This proves that the lens space is simultaneously $\mathbb{S}^3/\mathbb{Z}_m$ and the gluing of two solid tori, thus solving Exercise 5.3. As principal U_1 -bundles, these bundles are known as *monopole bundles*.

We thus have the surjective vertical morphism of principal bundles $\vartheta : \mathbb{S}^3 \rightarrow L(m; 1)$ given by the quotient by the \mathbb{Z}_m -action and the group homomorphism $\lambda : U_1 \rightarrow U_1$ given by the m -fold covering of the circle. Note that these bundles are also defined for m negative as the duals (inverse transition functions), or via the inverse U_1 -action on \mathbb{S}^3 .

15.3 Homogeneous bundles

We now consider the tangent bundle of a Klein geometry. Since it is a locally trivial principal bundle, we have local diffeomorphisms of the form $T(\pi^{-1}(U)) \cong T(U \times H) \rightarrow TU \times TH$. Local trivializations of TU and TH are given by the Maurer-Cartan forms Θ_G and Θ_H , respectively, so the diagram

$$\begin{array}{ccccc} T_g(gH) & \longrightarrow & T_gG & \longrightarrow & T_m(G/H) \\ \cong \downarrow \Theta_H & & \cong \downarrow \Theta_G & & \cong \downarrow \chi_g \\ \mathfrak{h} & \hookrightarrow & \mathfrak{g} & \longrightarrow & \mathfrak{g}/\mathfrak{h} \end{array} \quad (15.3.1)$$

is commutative with exact rows. Since \mathfrak{h} is only a subalgebra (not an ideal), the quotient $\mathfrak{g}/\mathfrak{h}$ is merely a vector space. The linear isomorphism χ_g is the unique map making the diagram commute.

We can identify the tangent space of G/H at m with $\mathfrak{g}/\mathfrak{h}$, but the identification χ_g depends on the choice of $g \in G$ over m . In fact, since $\pi \circ R_h = \pi$, the relation $R_h^* \Theta_H = \text{Ad}_h^{-1} \Theta_H$ implies that $\chi_{gh} = \text{Ad}_h^{-1} \circ \chi_g$. It follows that the identification of $T_m(G/H)$ with $\mathfrak{g}/\mathfrak{h}$ is determined only up to the adjoint action of H on $\mathfrak{g}/\mathfrak{h}$. This accounts for the frequent occurrence of the adjoint action in the sequel.

Proposition 15.3.1 ([41, Prop. 4.5.1]). $T(G/H) \cong G \times_H \mathfrak{g}/\mathfrak{h}$ as **VBs** over G/H .

Proof. Define a map $\varphi : G \times \mathfrak{g} \rightarrow T(G/H)$ by $\varphi(g, X) = (\pi(g), \pi_* \circ L_{g*}X)$, where $\pi : G \rightarrow G/H$ is the canonical projection. Clearly φ is smooth, surjective, and fiberwise linear. Moreover, for $X \in T_e H$, $L_{g*}X \in T_g(gH) \in \ker \pi_*$. Thus, $\varphi(g, X) = (g, 0)$. On the other hand,

$$\begin{aligned}\varphi(gh, \text{Ad}_h^{-1}X) &= (\pi(gh), \pi_*L_{gh*}\text{Ad}_h^{-1}X) = (\pi(g), \pi_*(L_{gh*}L_{h*}^{-1}R_{h*}X)) = \\ &= (\pi(g), L_{g*}\pi_*(R_{h*}X)) = (\pi(g), L_{g*}(\pi R_h)_*X) = (\pi(g), L_{g*}(\pi_*X)) = \varphi(g, X).\end{aligned}\quad (15.3.2)$$

Thus φ induces a smooth quotient map $\bar{\varphi} : G \times_H \mathfrak{g}/\mathfrak{h} \rightarrow T(G/H)$. Moreover, this map is injective since $\varphi(g, X) = \varphi(g', X')$ implies $g' = gh$ for some $h \in H$ and $\pi_*L_{gh*}X' = \pi_*L_gX$. The latter means that $\pi_*L_{h*}X' = \pi_*X$ and hence that $X' - \text{Ad}_h^{-1}X \in \mathfrak{h}$. Thus,

$$(g', X') = (gh, \text{Ad}_h^{-1}X \mod \mathfrak{h}) = (g, X) \text{ in } G \times_H \mathfrak{g}/\mathfrak{h}. \quad (15.3.3)$$

Thus $\bar{\varphi}$ is a vertical **VB** isomorphism. ■

- Example 15.3.2.**
1. Taking $\mathbb{S}^n = \text{SO}_{n+1}/\text{SO}_n$, we get the identification $T\mathbb{S}^n \cong \text{SO}_{n+1} \times_{\text{SO}_n} \mathbb{R}^n$, where \mathbb{R}^n is the standard (defining) representation of SO_n .
 2. Since the Grassmannian manifolds can be written as $G_k(\mathbb{K}^n) \cong O_n/(O_n - k \times O_k)$, by comparing transition functions it is easy to see that $TG_k(\mathbb{K}^n) \cong \text{Hom}(\gamma_k(\mathbb{K}^n), \gamma_k(\mathbb{K}^n)^\perp)$, where $\gamma_k(\mathbb{K}^n)$ is the tautological bundle over $G_k(\mathbb{K}^n)$ and $\gamma_k(\mathbb{K}^n)^\perp$ is its orthogonal complement in $G_k(\mathbb{K}^n) \times \mathbb{K}^n$.

Since tangent bundles involve only first derivatives, we can make the following rough division of Klein geometries into two types according to whether H is faithfully represented by the adjoint action on the tangent space.

Definition 15.3.3 (First order Klein geometry). A Klein geometry (G, H) is of *first order* if the representation $\text{Ad}_{\mathfrak{g}/\mathfrak{h}} : H \rightarrow \text{GL}(\mathfrak{g}/\mathfrak{h})$ is faithful. Otherwise, the geometry is said to have *higher order*.

The tangent bundle of a Klein geometry is an example of an important class of bundles that are equivariant with respect to the natural G -action \hat{L} on the base G/H in the following sense.

Definition 15.3.4 (Homogeneous bundle). Let G be a Lie group and $H \triangleleft G$ a closed subgroup. A (left) homogeneous, or *equivariant*, **FB** of type (G, H) is a **FB** $E \xrightarrow{\pi} G/H$ over the homogeneous space of (left) cosets G/H together with a (left) action $G \curvearrowright E$ that lifts the natural (left) action \hat{L} on G/H , i.e. which satisfies

$$\pi \circ \tilde{L}_g = \hat{L}_g \circ \pi, \text{ for all } g \in G. \quad (15.3.4)$$

In particular, a homogeneous **VB** is a **VB** which is also a homogeneous **FB** such that the G -action is by **VB** morphisms. Similarly, a homogeneous **PFB** is a **PFB** which is also a homogeneous **FB** such that the G -action is by **PFB** morphisms, i.e. by maps equivariant w.r.t. the (right) action of the structure group.

A morphism of homogeneous bundles over G/H is a vertical G -equivariant bundle morphism. The categories of homogeneous **FBS** and **VBS** are denoted $\text{Bun}_{(G,H)}$ and $\text{VectBun}_{(G,H)}$, respectively.

Example 15.3.5. The principal H -bundle $G \rightarrow G/H$ is obviously homogeneous because the left action L is a lift of \hat{L} .

One common way of producing homogeneous bundles is by applying a natural bundle functor F to G/H . Since by definition such a functor maps local diffeomorphisms to bundle morphisms, $F(\hat{L}_g)$ is a G -action as required in the definition of a homogeneous bundle. Thus, any natural bundle over a homogeneous space is homogeneous. As we will see below, the tangent bundle functor T that we considered in the previous section is as an example of this.

Let $E \rightarrow G/H$ be a homogeneous bundle. Consider the basepoint $[e] = eH \in G/H$ and the fiber $E_{[e]}$. The G -action on E restricts to an H -action on F , thus we can build the associated bundle $G \times_H E_{[e]} \rightarrow G/H$. We now show that this bundle is isomorphic to E , thus characterizing all homogeneous bundles as bundles associated to principal bundles of the form $G \rightarrow G/H$.

Theorem 15.3.6 ([42, Prop. 1.4.3]). *The mapping $E \rightarrow E_{[e]}$ (and restrictions of morphisms to the fiber) induces equivalences of categories $\text{Bun}(G, H) \simeq H\text{-Man}$ and $\text{VectBun}(G, H) \simeq H\text{-FinVect}$, where $H\text{-FinVect}$ is the category of finite-dimensional representations of H .*

Proof. Consider the map $G \times E_{[e]} \rightarrow E$ given by $(g, f) \mapsto g \times f$, where we use the action of G on E . Since the action $H \curvearrowright E_{[e]}$ is simply a restriction of this action, this map descends to a map $\varphi : G \times_H E_{[e]} \rightarrow E$ which is smooth since the projection $G \times E_{[e]} \rightarrow G \times_H E_{[e]}$ is a surjective submersion. Moreover, the definition $\varphi = g \cdot f$ implies that φ covers the identity and is also right G -equivariant, so it is a morphism of homogeneous bundles.

Now we build the inverse morphism. If $p \in E$, then by transitivity there exists a $g \in G$ such that $g \cdot \pi_E(p) = [e]$. Then $\pi_E(g \cdot p) = [e]$, so $g \cdot p \in E_{[e]}$, and we may consider $[(g^{-1}, g \cdot p)] \in G \times_H E_{[e]}$. This is independent of the choice of g , since for another choice $g' \in G$ we must have $(g'g^{-1}) \cdot p = p$, thus $g'g^{-1} \in H$, and thus

$$[(g'^{-1}, g' \cdot p)] = [(g'^{-1}g'g^{-1}, gg'^{-1}g' \cdot p)] = [(g^{-1}, g \cdot p)]. \quad (15.3.5)$$

Thus we get a well-defined map $\psi : E \rightarrow G \times_H E_{[e]}$. Choosing a local smooth section σ of $G \rightarrow G/H$ we can locally write $\psi(p) = [(\sigma(\pi_e(p))^{-1}, \sigma(\pi_E(p)) \cdot p)]$, which shows that ψ

is smooth. One immediately verifies that ψ and φ are mutually inverse bundle isomorphisms. To verify functoriality, if $\vartheta : E \rightarrow E'$ is a morphism of homogeneous bundles, then $\vartheta(E_{[e]}) \subset E'_{[e]}$ and the restriction $\vartheta|_{E_{[e]}} : E_{[e]} \rightarrow E'_{[e]}$ is H -equivariant. Equivariance of ϑ implies that for $g \in G$ and $f \in E_{[e]}$, we get $\vartheta(g \cdot f) = g \cdot \vartheta(f)$ which means that $\vartheta(\varphi([(g, f)])) = \varphi'([(g, \vartheta|_{E_{[e]}}(f))])$, so we see that identifying E and E' with $G \times_H E_{[e]}$, respectively, $G \times_H E'_{[e]}$, the map ϑ is induced by $\text{id} \times \vartheta|_{E_{[e]}}$. Conversely, any H -equivariant map $F \rightarrow F'$ clearly induces a morphism of associated homogeneous bundles.

Finally, note that for a homogeneous **VB**, the above construction clearly produces an isomorphism $G \times_H E_{[e]} \rightarrow E$ of homogeneous **VBS**, and homogeneous **VBS**s correspond to linear H -equivariant maps between their typical fibers. ■

Thus an alternative definition of homogeneous bundles is simply as bundles associated to principal bundles of the form $G \rightarrow G/H$. Functoriality implies that these equivalences are compatible with various constructions such as fibered products, Whitney sums, tensor products, etc. Conversely, for example, a decomposition of a tensor product of representations into a direct sum of smaller representations induces a similar decomposition of corresponding Whitney products of homogeneous **VBS**s.

Example 15.3.7. The tangent bundle $T(G/H)$ that we studied in Proposition 15.3.1 is clearly homogeneous because the left action \hat{L}_g on G/H induces a lifted action \hat{L}_{g*} on $T(G/H)$, and it corresponds to the H -representation on $\mathfrak{g}/\mathfrak{h}$ coming from the restriction to H of the adjoint representation of G . The isomorphism with $G \times_H (\mathfrak{g}/\mathfrak{h})$ established in that proposition is exactly the same as what we would obtain from Proposition 15.3.6.

By the naturality of the correspondence, the cotangent bundle $T^*(G/H)$ corresponds to the dual representation $(\mathfrak{g}/\mathfrak{h})^*$. Note that the latter is just the annihilator of \mathfrak{h} in \mathfrak{g}^* , so this is a subrepresentation of the restriction of the coadjoint representation of G to H . Again, by naturality, the tensor bundle $T_l^k G/H$ corresponds to $T_l^k(\mathfrak{g}/\mathfrak{h}) = (\mathfrak{g}/\mathfrak{h})^{\otimes k} \otimes (\mathfrak{g}/\mathfrak{h})^{*\otimes l}$.

Finally we provide a similar classification of principal homogeneous bundles.

Proposition 15.3.8 ([42, Lem. 1.4.5]). *Let (G, H) be a homogeneous space and K a Lie group. Let $P \rightarrow G/H$ be a homogeneous principal K -bundle of type (G, H) . Then there is a smooth homomorphism $\lambda : H \rightarrow K$ such that $P \cong G \times_H K = G^{[\lambda]}$, where the action of H on K is $h \cdot k = \lambda(h)k$.*

The bundles corresponding to two homomorphisms $\lambda, \lambda' : H \rightarrow K$ are (vertically) isomorphic iff λ and λ' are conjugate, i.e. $\lambda'(h) = k\lambda(h)k^{-1}$ for some $k \in K$ and all $h \in H$.

Proof. Let $P_{[e]}$ be the fiber of P over $[e] = eH \in G/H$. The left G -action on P restricts to a left H -action on $P_{[e]}$, and $P \cong G \times_H P_{[e]}$ as homogeneous bundles. Fixing a point

$p_0 \in P_{[e]}$, the map $k \mapsto p_0 \cdot g$ is a diffeomorphism $K \rightarrow P_{[e]}$, so it remains to describe the H -action in this picture.

For $h \in H$, we have $h \cdot p_0 \in P_{[e]}$, so there is a unique element $\lambda(h) \in K$ such that $h \cdot p_0 = p_0 \cdot \lambda(h)$. By smoothness of the two actions, the map $\lambda : H \rightarrow K$ defined this way is smooth, and $\lambda(e) = e$. Since H acts by principal bundle maps, we see that $h \cdot (p_0 \cdot k) = (h \cdot p_0) \cdot k = p_0 \cdot \lambda(h)k$. From here, λ is a homomorphism.

Suppose that we are given an isomorphism $G \times_{\lambda} K \rightarrow G \times_{\lambda'} K$ of homogeneous PFBs/ Then the restriction to the fibers over $[e]$ is a diffeomorphism $\phi : K \rightarrow K$ which commutes with the principal right K -action and is equivariant for the two left H -actions. By the first property, $\phi(k) = k_0 k$, where $k_0 = \phi(e)$. But then the second property reads as follows:

$$k_0 \lambda(h)k = \phi(\lambda(h)k) = \lambda'(h)\phi(k) = \lambda'(h)k_0 k. \quad (15.3.6)$$

In particular, $\lambda'(h) = k_0 \lambda(h)k_0^{-1}$. Conversely, if λ and λ' are conjugate, right multiplication by k_0 induces a diffeomorphism with the two required equivariance properties. From Proposition 15.3.6 we conclude that this gives rise to an isomorphism of homogeneous principal bundles. ■

15.4 Frobenius reciprocity

Consider the set $\Gamma(E)$ of sections of a homogeneous vector bundle $E \rightarrow G/H$. There is a natural left G -action on $\Gamma(E)$ defined by

$$g \cdot \sigma := \tilde{L}_g \circ \sigma \circ \hat{L}_g^{-1}, \quad (15.4.1)$$

where \tilde{L} and \hat{L} are the actions on E and G/H , respectively. Note that this particular combination of the two actions is required so that sections get mapped to sections, i.e. so that the action preserves the point in the base.

Definition 15.4.1 (Induced representation). Let (G, H) be a homogeneous space, $G \curvearrowright V$ a finite-dimensional linear representation G , and $E = G \times_H V$ be the corresponding homogeneous VB. The representation of G on $\Gamma(E)$ defined above is called the *induced representation* of G , denoted $\text{Ind}_H^G(V)$. The association functor for bundles makes this a functor of V .

In addition, we denote by $\text{Res}_H^G(V)$ the representation of H on G obtained by restriction of the representation of G . This is also obviously functorial in V .

The correspondence between sections of an associated bundle and equivariant functions on the principal bundle (Proposition 14.2.17) gives rise to another interpretation of the induced action of G on sections. This picture is (in the case of homogeneous VBs) frequently used in representation theory. Namely, via the isomorphism $E \cong G \times_H E_{[e]}$ from Proposition 15.3.6,

we can identify $\Gamma(E)$ with $\Gamma(G \times_H E_{[e]})$, which in turn by Proposition 14.2.17 is in bijective correspondence with the set $\text{Hom}_H(G, E_{[e]})$. Explicitly, the correspondence between a section σ and a function φ is characterized by $\sigma(gH) = [(g, \varphi(g))] = g \cdot \varphi(g)$ or by $\varphi(g) = g^{-1} \cdot \sigma(gH)$. In the case of a homogeneous vector bundle, this bijection is clearly a linear isomorphism.

To get the action of G on $\Gamma(E)$ in the picture of equivariant functions, we only have to notice that by definition $(g \cdot \sigma)(g'H) = g \cdot \sigma(g^{-1}g'H)$. Consequently, the map $g \cdot f : G \rightarrow E_{[e]}$ corresponding to $g \cdot \sigma \in \Gamma(E)$ is given by none other than the natural action by pullbacks L_g^* :

$$(g \cdot \varphi)(g') = g'^{-1} \cdot (g \cdot \sigma(g^{-1}g'H)) = (g^{-1}g')^{-1} \cdot \sigma(g^{-1}g'H) = \varphi(g^{-1}g'). \quad (15.4.2)$$

The next fundamental result shows that in certain situations one can reduce questions about the infinite-dimensional space $\Gamma(E)$ to questions about the finite-dimensional manifold $E_{[e]}$. In particular, determining G -invariant sections of a homogeneous bundle always reduces to a finite-dimensional problem.

Theorem 15.4.2 (Geometric Frobenius reciprocity [42, Thm. 1.4.4]). *Let $E \rightarrow G/H$ be a homogeneous bundle with standard fiber $E_{[e]}$ (viewed as an H -space), and let X be a smooth manifold endowed with a smooth left G -action. Then the evaluation at $[e]$ induces a natural bijection between the set of G -equivariant maps $X \rightarrow \Gamma(E)$ and the set of H -equivariant maps $X \rightarrow E_{[e]}$. In particular, there is a natural bijection between the set $\Gamma(E)^G$ of G -invariant sections of E and the set $(E_{[e]})^H$ of H -invariant elements of the typical fiber. If E is the homogeneous VB induced by an H -representation W , and V is a representation of G , then the bijections above are linear and respect the subspaces of linear equivariant maps, so that there is a natural linear isomorphism*

$$\text{Hom}_G(V, \text{Ind}_H^G(W)) \cong \text{Hom}_H(\text{Res}_H^G(V), W), \quad (15.4.3)$$

i.e. Res_H^G and Ind_H^G are adjoint functors.

Proof. Suppose are given a map $X \rightarrow \Gamma(E)$ written as $x \mapsto \sigma_x$. This map is G -equivariant iff $\sigma_{g \cdot x}(g'H) = g \cdot \sigma_x(g^{-1}h'H)$. In particular, if we consider the map $X \rightarrow E_{[e]}$ defined by $x \mapsto \sigma_x([e])$, then for $h \in H$, we get $\sigma_{h \cdot x}([e]) = h \cdot \sigma_x(h^{-1} \cdot [e]) = h \cdot \sigma_x([e])$, so this is H -equivariant.

Conversely, if $\varphi : X \rightarrow E_{[e]}$ is H -equivariant, then for $x \in X$ we define $\sigma_x : G/H \rightarrow E$ by $\sigma_x(gH) = g \cdot \varphi(g^{-1} \cdot x)$. This is well defined since $\varphi(h^{-1}g^{-1} \cdot x) = h^{-1} \cdot \varphi(g^{-1} \cdot x)$, and thus $gh \cdot \varphi(h^{-1}g^{-1} \cdot x) = g \cdot \varphi(g^{-1} \cdot x)$. Moreover, $\sigma_x([e]) = \varphi(x)$ and

$$\sigma_{g \cdot x}(g'H) = g' \cdot \varphi(g'^{-1}g \cdot x) = g \cdot g^{-1}g' \cdot \varphi(g'^{-1}g \cdot x) = g \cdot \sigma_x(g^{-1}g'H), \quad (15.4.4)$$

so $x \mapsto \sigma_x$ is G -equivariant. Since for any G -equivariant map $x \mapsto \sigma_x$ we must have $\sigma_x(gH) = g \cdot g^{-1} \cdot \sigma_x(g^{-1}gH) = g \cdot \sigma_{g \cdot x}([e])$, the two constructions are inverse bijections between the set of G -equivariant maps $X \rightarrow \Gamma(E)$ and the set of H -equivariant maps $X \rightarrow E_{[e]}$.

Taking $X = \{*\}$ a single point with the trivial G -action, a G -equivariant map $X \rightarrow \Gamma(E)$ is just a G -invariant section $g \cdot \sigma_* = \sigma_*$, and a H -equivariant map $X \rightarrow E_{[e]}$ is an H -invariant element in $E_{[e]}$. Thus we get a bijection $\Gamma(E)^G \rightarrow (E_{[e]})^H$.

Finally, if E is a natural **VB** induced by an H -representation W and X is a G -representation V , then $L(V, \Gamma(E))$ and $L(V, E_{[e]})$ are vector spaces under the pointwise operations, and the evaluation at $[e]$ induces a linear map $L(V, \Gamma(E)) \rightarrow L(V, E_{[e]})$. If we start from a linear map $\varphi : V \rightarrow E_{[e]}$, the construction above produces $\sigma_v(gH) = g \cdot \varphi(g^{-1} \cdot v)$. Since φ is linear and G acts by linear maps, this is linear in v . ■

Example 15.4.3 (Spherical harmonics). Looking ahead, we can take H to be a maximal torus subgroup $T \triangleleft G$. Since T is abelian, its irreducible representations (irreps) are one-dimensional and are called *weights*, labeled by λ . The associated homogeneous line bundles over G/T are denoted L_λ . Further if V_γ is an irrep of G (labeled by a highest weight γ , irrelevant to this statement), then by Schur's Lemma the dimension of the space $\text{Hom}_G(V_k, \Gamma(L_n))$ is simply counting the multiplicity of V_k in the decomposition of $\Gamma(L_n)$ into irreps. Therefore the Frobenius reciprocity in the language of representation theory states that the multiplicity of a G -irrep V_γ in the decomposition of $\Gamma(L_n)$ is equal to the multiplicity of the weight λ in V_γ . If G is compact, then for any given weight λ , the induced representation $\Gamma(L_\lambda)$ will decompose into an infinity of irreducibles (because each is finite-dimensional for compact G), namely all those that contain the weight λ .

Consider $G = \text{SU}_2$ and $T = \text{U}_1$ the subgroup fixing a vector. Then $G/T \cong \mathbb{S}^2$. Since T is abelian, all of its representations decompose into sums of one-dimensional (hence irreducible) ones, and the one-dimensional representations are labeled by an integer $n \in \mathbb{Z}$ (namely the representation is given by multiplication by z^n , where $z \in \text{U}_1 \subset \mathbb{C}$). Meanwhile the irreducible representations V_k of SU_2 are known to be labeled by non-negative integers k , where k is the largest integer that occurs when you restrict V_k to be a U_1 representation. Frobenius reciprocity implies

$$\dim \text{Hom}_{\text{SU}_2}(V_k, \Gamma(L_n)) = \dim \text{Hom}_{\text{U}_1}(V_k, n) \quad (15.4.5)$$

where L_n is the complex line bundle over \mathbb{S}^2 corresponding to n . The right hand side is the multiplicity of the weight n in V_k (that is, the multiplicity of n in the restriction of the representation on V_k to T). Since we know what the weights of V_k are ($-k, -k+2, \dots, k-2, k$, all with multiplicity one), we can see that the decomposition of $\Gamma(L_n)$ into irreducibles will include all V_k where k is the same parity as n and $k \geq n$.

In the case $n = 0$, L_0 is the trivial \mathbb{C} -bundle and $\Gamma(E_0) \cong C^\infty(\mathbb{S}^2, \mathbb{C})$ is just the space of complex-valued functions on \mathbb{S}^2 . Only V_k with even k include the weight 0, therefore $C^\infty(\mathbb{S}^2, \mathbb{C})$ decomposes into a sum of such irreps. To compute the multiplicity of each of them, we use Frobenius reciprocity. As stated above, the multiplicity of $n = 0$ in V_{2l} is 1. For analytic reasons that will be clarified in the Peter-Weyl theorem, the direct sum of all V_{2l} 's, once endowed with a proper topology, is not exactly $C^\infty(\mathbb{S}^2)$, but rather the Hilbert space $L^2(\mathbb{S}^2)$ of square integrable functions:

$$L^2(\mathbb{S}^2, \mathbb{C}) \cong \bigoplus_{l=0}^{\infty} V_{2l}. \quad (15.4.6)$$

This holds both over reals and complexes. The space V_{2l} is $(2l + 1)$ -dimensional and a standard basis is given by the spherical harmonics $Y_m^l(\theta, \varphi)$. If we replace SU_2 by SO_3 , the only difference is that V_k with odd k disappear from the list of irreps, so the final result for $n = 0$ is the same. This is the beginning of the subject of *harmonic analysis* (and in particular Fourier transforms) on Lie groups.

The geometric version of Frobenius reciprocity immediately allows us to reduce questions about the existence of invariant Riemannian metrics and other invariant tensor fields to problems of finite-dimensional representation theory: if $M = G/H$ is a homogeneous space, then from Proposition 15.3.6 we know that the tensor bundle $T_l^k M$ is associated to $G \rightarrow M$ with respect to the representation $T_l^k(\mathfrak{g}/\mathfrak{h})$. By Frobenius reciprocity, G -invariant sections of this bundle are in bijective correspondence with H -invariant elements in the representation. Since invariant elements are the same as homomorphisms from the trivial representation, this can be rephrased in such a way that the dimension of invariant tensor fields of type $\binom{k}{l}$ on G/H equals the multiplicity of the trivial representation in $T_l^k(\mathfrak{g}/\mathfrak{h})$. Moreover, the proof of Theorem 15.4.2 gives an explicit construction of the invariant tensor field from the invariant element in the representation. Clearly, the whole construction respects symmetries of any kind, so invariant tensor fields having certain symmetries are in correspondence with invariant elements of the representations having the same symmetries. Finally, if we consider (pointwise) questions of non-degeneracy they clearly reduce to the analogous non-degeneracy properties in the representation. In particular, the set of G -invariant Riemannian metrics on G/H is in bijective correspondence with the set of H -invariant scalar products on $\mathfrak{g}/\mathfrak{h}$. Hence, we get

Corollary 15.4.4. *A homogeneous space G/H admits a G -invariant Riemannian metric iff the image $H_1 \subset GL(\mathfrak{g}/\mathfrak{h})$ of H under the map induced by the restriction to H of the adjoint representation of G has compact closure in $GL(\mathfrak{g}/\mathfrak{h})$.*

Proof. From above we know that existence of an invariant Riemannian metric is equivalent to existence of an H_1 -invariant positive definite scalar product on $\mathfrak{g}/\mathfrak{h}$. If such a scalar product exists, then H_1 is contained in the orthogonal group of this product, which is compact. Hence H_1 has compact closure. Conversely, if $K \subset \mathrm{GL}(\mathfrak{g}/\mathfrak{h})$ is a compact subgroup containing H_1 , then averaging any scalar product on $\mathfrak{g}/\mathfrak{h}$ over K (as in Proposition 11.2.8) gives a K -invariant and thus a H_1 -invariant scalar product. ■

- Example 15.4.5.**
1. Viewing the sphere \mathbb{S}^n as the homogeneous space O_{n+1}/O_n , we have $H = O_n$ and the H -representation $\mathfrak{g}/\mathfrak{h} \cong \mathbb{R}^n$ is just the standard one, so the set of O_{n+1} -invariant Riemannian metrics on \mathbb{S}^n consists exactly of all constant positive multiples of the standard metric (since the same statement holds on \mathbb{R}^{n+1}).
 2. The description of invariant Riemannian metrics on \mathbb{R}^n , viewed as a homogeneous space of the Euclidean group E_n is exactly the same.
 3. For the projective sphere we have $H_1 = \mathrm{GL}(\mathfrak{g}/\mathfrak{h})$, so there is certainly no $\mathrm{SL}_{n+1}(\mathbb{R})$ -invariant metric on \mathbb{S}^n . Similarly, one shows that on the contact, conformal, and CR spheres there are no invariant Riemannian metrics.

15.5 Local gauges

For each Klein geometry (G, H) we have the corresponding pair of Lie algebras $(\mathfrak{g}, \mathfrak{h})$. If (G, H) is effective, then \mathfrak{h} contains no nontrivial ideals of \mathfrak{g} .

Definition 15.5.1 (Infinitesimal Klein geometry). An infinitesimal Klein geometry, or a Klein pair, is a pair of Lie algebras $(\mathfrak{g}, \mathfrak{h})$ where \mathfrak{h} is a subalgebra of \mathfrak{g} . Its *kernel* \mathfrak{k} is the largest ideal of \mathfrak{g} contained in \mathfrak{h} . If $\mathfrak{k} = 0$, we call the pair *effective*. If there is an \mathfrak{h} -module decomposition $\mathfrak{h} = \mathfrak{h} \oplus \mathfrak{m}$, we call it *reductive*.

Not every Klein pair can be realized as the pair of a Klein geometry.

Theorem 15.5.2 ([41, Thm. 4.1]). *Let (G, H) be a Klein geometry with kernel K and associated Klein pair $(\mathfrak{g}, \mathfrak{h})$ with kernel \mathfrak{k} . Then*

$$K = \{h \in H \mid \mathrm{Ad}_h X - X \in \mathfrak{k} \text{ for all } X \in \mathfrak{g}\}. \quad (15.5.1)$$

Corollary 15.5.3 (Fundamental property of effective Klein geometries [41, Cor. 4.2]). *If (G, H) is an effective Klein geometry, N is a subgroup of H , and $\mathfrak{n} = \mathrm{Lie} N$, then*

$$N = \{h \in H \mid \mathrm{Ad}_h X - X \in \mathfrak{n} \text{ for all } X \in \mathfrak{g}\} \Rightarrow N = \{e\}. \quad (15.5.2)$$

We now provide a local interpretation of Klein geometry, starting with an abstract example.

Example 15.5.4 (Meteor tracking). Suppose we live in a small open set $U \subset M = G/H$ and a meteor flashes through U . We wish to describe its motion. We formalize the meteor as a point set X , and a *configuration* of the meteor is a map $X_0 : X \rightarrow M$. This is the origin for Cartan's idea of a "moving frame".

We assume the meteor is *rigid* and *sufficiently complicated*. By the term rigid in the given geometry, we mean that for any two of its positions, where it has the configurations X_0 and X_1 say, there is an element of G carrying X_0 to X_1 . By sufficiently complicated we mean that the subgroup of G that fixes the body pointwise (its stabilizer) is the identity. It follows that if $X(t)$ is the configuration of its points at time t , then there is a unique path $g(t) \in G$ such that $X(t) = g(t)X(0)$. One way to describe the motion would be to specify the path $g(t)$ itself. However, it is more useful, and closer to what is actually observed, to describe the motion differently. We first describe the motion of one of its points $q \in X(0)$, which we take to be eH , by a path $q(t) = g(t)H$ on U , and then describe the motion of the rest of the body as turning about this one point as it moves along. For this we will need the concept of a gauge. Let us assume that U is small enough to admit a local section $\sigma : U \rightarrow G$. Any two such sections σ, σ' are related by a map $h : U \rightarrow H$ by $\sigma'(u) = \sigma(u)h(u)$. A "choice of gauge" is a choice of such a local section, and $h(u)$ is called a gauge transformation. A choice of gauge can be regarded as a choice of motions, varying smoothly with u , which maps the base point to u . The value of a gauge at any point may be called a *frame* at that point and the gauge itself may be called a *moving frame*. From this point of view the principal bundle $G \rightarrow G/H$ is a bundle of frames.

Since both $\sigma(q(t))$ and $g(t)$ lie over $q(t)$, they must differ by an element $k(t) \in H$. Thus, in the presence of a gauge, the motion can be described by giving $q(t)$ and $(k(t))$. The latter describes the way the meteor turns about the point $q(t)$. Crucially, there is no intrinsic geometric meaning attached to the choice of gauge σ , or the function $k(t)$ alone. Together, however, they complete the description of the motion.

Proposition 15.5.5 ([41, Prop. 4.7.1]). *If $\chi : \pi^{-1}(U) \rightarrow U \times H$ is a local chart for the principal bundle $G \rightarrow G/H$, then the local representative of the Maurer-Cartan form Θ_G in this chart at point (u, h) is $\text{Ad}_h^{-1}\Theta_G \oplus \Theta_H$, that is,*

$$(\chi^{-1})^*\Theta_G(X_u, Y_h) = \text{Ad}_h^{-1}\Theta_G(X_u) + \Theta_H(Y_h). \quad (15.5.3)$$

Proof. Let $\sigma : U \rightarrow G$ be the local section defining χ , i.e. $\sigma(u) = \chi^{-1}(u, e)$. Also let $i_h : U \rightarrow U \times H$ and $j_u : H \rightarrow U \times H$ be the inclusions of slices $u \mapsto (u, h)$ and $h \mapsto (u, h)$, respectively. Then we have

$$(\chi^{-1})^*\Theta_G \circ i_{h*} = (\chi^{-1} \circ i_h)^*\Theta_G = (R_h\sigma)^*\Theta_G = \sigma^*R_h^*\Theta_G = \sigma^*(\text{Ad}_h^{-1}\Theta_G) = \text{Ad}_h^{-1}\sigma^*\Theta_G, \quad (15.5.4)$$

and on the other hand,

$$(\chi^{-1*}\Theta_G) \circ j_{u*} = (\chi^{-1} \circ j_u)^*\Theta_G = L_{\sigma(u)}^* \Theta_G|_{T_h H} = \Theta_G|_{T_h H} = \Theta_H. \quad (15.5.5)$$

■

Let $\sigma : U \rightarrow G$ be a local gauge, that is, a local section of the principal H -bundle $G \rightarrow G/H$ over an open set $U \subset G/H$. Then σ pulls back the Maurer-Cartan form Θ_G to a \mathfrak{g} -valued 1-form θ_U on U : $\sigma^*\Theta_G = \theta_U$. The structural (Maurer-Cartan) equation for Θ_G then yields

$$d\theta_U + \frac{1}{2}[\theta_U, \theta_U] = 0. \quad (15.5.6)$$

Note that θ_U *almost* determines σ since the Fundamental Theorem 11.3.2 says that θ_U determines σ up to left multiplication by a fixed element of G . Indeed, for an effective Klein geometry, if σ is a section, then $g\sigma$ is a section only if $g = e$.³⁴ For this reason we may refer to θ_U as an *infinitesimal gauge*. We have the non-canonical isomorphism at every $u \in U$

$$\bar{\theta}_U : T_u U \xrightarrow{\theta_U} \mathfrak{g} \xrightarrow{\text{mod } \mathfrak{h}} \mathfrak{g}/\mathfrak{h}. \quad (15.5.7)$$

The infinitesimal gauge θ_U can be roughly interpreted as assigning to each tangent vector $X \in T_u M$ the infinitesimal motion $I + \epsilon\theta(X) \in G$ of M whose effect on u itself is to move it to $u + \epsilon X$.

By differentiating the transformation rule $\sigma'(u) = \sigma(u)h(u)$ we get

$$\sigma'^*\Theta_G = \text{Ad}_h^{-1}\sigma^*\Theta_G + h^*\Theta_H, \quad (15.5.8)$$

that is

$$\theta'_U = \text{Ad}_h^{-1}\theta_U + h^*\Theta_H. \quad (15.5.9)$$

This transformation is called an infinitesimal change of gauge, and the two infinitesimal gauges are called gauge equivalent. We denote this symbolically by $\theta_U \Rightarrow_h \theta'_U$.

16 Connections

16.1 General connections

Connections are a large and varied topic. Our exposition is based on a combination of [44, 17, 43].

The Frobenius Theorem 9.5.4 states that a distribution $\mathcal{H} \subset TM$ is integrable iff it is involutive. However, there is no clear *quantitative measure* of integrability. It turns out that one way to obtain such a quantitative measure is to assume that there is a second distribution, *transverse* to the first one. Namely, let $(\mathcal{H}, \mathcal{V})$ be two distributions on M (subbundles of TM) such that

$$\mathcal{H}_m \oplus \mathcal{V}_m = T_p M \quad (16.1.1)$$

for all $m \in M$. This implies that there are globally defined smooth bundle maps (projections), which we denote by the same symbols,

$$\mathcal{H} : TM \rightarrow \mathcal{H}, \quad \mathcal{V} : TM \rightarrow \mathcal{V}, \quad \mathcal{H} + \mathcal{V} = \text{id}_{TM}. \quad (16.1.2)$$

Now we can construct a linear map that measures to what extent H fails to be an integrable distribution. Since \mathcal{H} and \mathcal{V} are completely interchangeable at this point, we in fact get two maps (the sign choice is conventional):

$$\mathcal{R} : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \Gamma(V), \quad \mathcal{R}(X, Y) = -\mathcal{V}[\mathcal{H}X, \mathcal{H}Y], \quad (16.1.3)$$

$$\bar{\mathcal{R}} : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \Gamma(H), \quad \bar{\mathcal{R}}(X, Y) = -\mathcal{H}[\mathcal{V}X, \mathcal{V}Y]. \quad (16.1.4)$$

Indeed, if $\mathcal{R} = 0$, it means that the Lie bracket of two sections of \mathcal{H} lies in \mathcal{H} , and therefore \mathcal{H} is integrable, and similarly $\bar{\mathcal{R}}$ detects the integrability of \mathcal{V} . It is obvious that these maps are $C^\infty(M)$ -linear and hence are tensors. We call \mathcal{R} *curvature* and $\bar{\mathcal{R}}$ *cocurvature*. Note that the value of \mathcal{R} depends on the choice of \mathcal{V} , but whether it is zero or not doesn't, so in this sense it's still an objective measure of integrability for \mathcal{H} (as long as \mathcal{H} and \mathcal{V} are everywhere transverse). It turns out that fiber bundles are special in that they come with one canonically defined integrable distribution that we define now. A more sophisticated way of defining the curvature of general connections is as the Fölicher-Nijenhuis bracket $[\mathcal{V}, \mathcal{V}]$, see [43, Sec. 9].

Definition 16.1.1 (Vertical bundle). If $E \xrightarrow{\pi} M$ is a smooth FB, then its *vertical subbundle* is the subbundle $\mathcal{V}E < TE$ defined as

$$\mathcal{V}E = \ker \pi_* \quad (16.1.5)$$

(it is a subbundle by virtue of being the kernel of a bundle morphism). At a point $p \in E$

the fiber is denoted $\mathcal{V}_p E$. Since each fiber E_m , $m \in M$, is an embedded submanifold, and $\mathcal{V}E$ restricted to E_m coincides with the tangent bundle TE_m , we conclude that $\mathcal{V}E$ is an integrable distribution. In particular, it is involutive: for any two *vertical vector fields* $X, Y \in \mathfrak{X}_{\text{ver}}(E) := \Gamma(\mathcal{V}E)$, we have $[X, Y] \in \mathfrak{X}_{\text{ver}}(E)$.

Therefore cocurvature on fiber bundles always vanishes. Crucially, there is no *natural* “horizontal” subbundle $\mathcal{H}E < TE$ such that $\mathcal{H}E \oplus \mathcal{V}E = TE$. A connection is, in the broadest sense, a choice of a transverse “horizontal bundle” $\mathcal{H}E$, and its integrability is quantified by the curvature tensor. If it is integrable, then the fiber bundle is not only foliated by its fibers, but also by “horizontal” integral submanifolds (locally horizontal sections).

Note that for $m \in M$ and $p \in E$, the sequence

$$0 \rightarrow T_p E_m \xrightarrow{T_p i_m} T_p E \xrightarrow{T_p \pi} T_m M \rightarrow 0 \quad (16.1.6)$$

is a short exact sequence. Here $m = \pi(p)$ and $E_m \xhookrightarrow{i_m} E$ is the natural inclusion. This sequence means that we can identify the quotient space $T_p E / \mathcal{V}_p E$ with $T_m M$, so we have the following *natural* isomorphism of **VBs** over E :

$$TE / \mathcal{V}E \cong \pi^* TM. \quad (16.1.7)$$

But first we will approach this problem from another angle.

Determining what the “horizontal” directions in a bundle are, in particular, provides one with a *unique* way of lifting paths γ in M to paths $\tilde{\gamma}$ in E . Doing this for every starting point $p \in E_{\gamma(0)}$, we get a map $E_{\gamma(0)} \rightarrow E_{\gamma(1)}$. This map has to be bijective because the lifting clearly respects inverses of paths. Moreover, it will depend only on the *shape* of γ and not on how exactly it is parametrized. The next three definitions formalize the concept of *reparametrization invariance* of paths.

Definition 16.1.2 (Sitting instants). A path $\gamma : I \rightarrow M$ in a topological space M is said to have sitting instants if there is a neighborhood of the points $0, 1$ in I on which γ is locally constant.

Definition 16.1.3 (Thin homotopy). A thin homotopy between two smooth paths $\gamma_0, \gamma_1 : I \rightarrow M$ in a smooth manifold M is a smooth homotopy $H : I^2 \rightarrow M$ between them such that the rank of its derivative $H_* : T(I^2) \rightarrow TM$ is everywhere at most 1. Equivalently, for any 2-form $\omega \in \Omega^2(M)$, its pullback must vanish: $H^* \omega = 0$ on I^2 .

Definition 16.1.4 (Path groupoid). The path groupoid $P_1(M)$ of a smooth manifold M is the groupoid whose objects are points of M and morphisms $\text{Mor}(p, q)$ are the thin homotopy classes of paths $p \sim q$ with sitting instants. Composition of paths is defined via concatenation and reparametrization as usual. The quotient by thin homotopies ensures that this is an

| associative product with inverses for each path.

This groupoid can in fact be understood as smooth in a certain sense (it is a “diffeological groupoid”), but we will not go into those details. We can now define a connection as a prescription (a functor) for taking a path γ in M and producing a diffeomorphism $E_{\gamma(0)} \rightarrow E_{\gamma(1)}$ in such a way that it is independent of reparametrizations of γ and consistent, i.e. functorial with respect to products of paths. In addition, it needs to depend on γ smoothly in a proper sense. Finally, if the bundle carries a G -structure, we would like this diffeomorphism to be a morphism of G -fibers, i.e. be represented in local frames by actions of elements of G on the typical fiber.

Definition 16.1.5 (Connection I: transport functor). Let $(E \xrightarrow{\pi} M, G \curvearrowright F, \mathcal{G})$ be a G -bundle, and more specifically an \mathcal{S} -bundle for some category \mathcal{S} . Let $\mathcal{S}\text{-Man}^\infty$ be the category of \mathcal{S} -fibers (Definition 7.3.1). A *transport functor* on E is a functor

$$\text{tra} : \mathbf{P}_1(M) \rightarrow \mathcal{S}\text{-Man}^\infty, \quad (16.1.8)$$

such that for any path $\gamma \in \mathbf{P}_1(M)$ the map $\text{tra}_\gamma := \text{tra}(\gamma)$ is a morphism (hence isomorphism) of the G -fibers over the endpoints:

$$\text{tra}_\gamma : (E_{\gamma(0)}, \mathcal{G}_{\gamma(0)}) \rightarrow (E_{\gamma(1)}, \mathcal{G}_{\gamma(1)}), \quad (16.1.9)$$

and which is smooth in the following sense. If γ_t is a smooth homotopy of paths in M (not necessarily with fixed ends), then, for any $t \in I$, in some local \mathcal{G} -charts of E at $\gamma_t(0)$ and $\gamma_t(1)$ the representatives of the diffeomorphism $\text{tra}_{\gamma_t} : E_{\gamma_t(0)} \rightarrow E_{\gamma_t(1)}$ must depend smoothly on t as elements of $\mathcal{S}(F, F)$. Since any isomorphism of G -fibers is represented by G -actions in G -frames, this can also be stated as there existing a (finite) collection of \mathcal{G} -charts covering the curves traced out by the endpoints $\gamma_t(0)$ and $\gamma_t(1)$ and a smooth path $g : I \rightarrow G$ such that the local representatives of tra_{γ_t} in any of those charts are exactly the action of $g(t)$ on F .

A *morphism of connections* on E is a natural transformation of two such functors.

Note that in the case of principal bundles the complicated condition on the values of tra can be replaced by simply saying that it is a functor into the category $G\text{-Tor}$ of G -torsors. This definition can be made significantly neater by introducing a few extra functors, but we direct the reader to the work of Schreiber and Waldorf (2014) for that.

Definition 16.1.6 (Flat connection). A connection defined via a transport functor tra on a bundle E is called flat if the value of tra_γ is independent of the homotopy class of γ with fixed endpoints.

Example 16.1.7 (Absolute parallelism). If $E \rightarrow M$ is a trivial **FB**, then a particular choice of a global trivialization $\chi : E \rightarrow M \times F$ is called an *absolute parallelism*. Each absolute parallelism determines a unique flat connection. Namely, the parallel transport operator (independent of the path) from fiber E_m to fiber $E_{m'}$ maps $\chi^{-1}(m, f)$ to $\chi^{-1}(m, f')$.

Example 16.1.8 (Canonical flat connections on a Lie group). If G is a Lie group, then the tangent bundle TG admits two obvious flat connections. The first, called the $(-)$ -connection, has parallel transport acting by left translations, i.e. $T_g G$ is mapped to $T_h G$ by $L_{gh^{-1}*}$. The second, so called $(+)$ -connection, has parallel transport given by $L_{gh^{-1}*}$ instead. Since these transports are independent of the path, these connections are flat. Both of them derive from absolute parallelisms $TG \rightarrow G \times \mathfrak{g}$ given by the left and right Maurer-Cartan forms, respectively.

Proposition 16.1.9. *A bundle is flat iff it admits a flat connection.*

Proof. If a bundle is flat, then by definition it has a G -atlas with locally constant transition functions. Then a path γ in the base can be turned into a “horizontal” path $t \mapsto (\gamma(t), f_\alpha)$ in each local chart $U_\alpha \times F$ overlapping with the image of γ . The points f_α are chosen so that $f_\beta = g_{\beta\alpha} \cdot f_\alpha$, where $g_{\alpha\beta} \in G$ is the constant transition function on $U_{\alpha\beta}$. This produces a consistently defined path in the bundle because transition functions are constant. By applying this to every starting point, we get a transport map which obviously depends only on the endpoints of γ , thus defining a flat connection.

Conversely, given a flat connection tra , we can cover the base manifold with contractible open sets U_α , pick a basepoint x_α in each, and use parallel transport along the paths $x_\alpha \sim y_\alpha$ inside U_α to obtain diffeomorphisms $\varphi_{\alpha, y_\alpha} : E_{y_\alpha} \rightarrow E_{x_\alpha}$ for all $y_\alpha \in U_\alpha$. Due to flatness these diffeomorphisms depend only on y_α . Hence we can define a map $\pi^{-1}(U_\alpha) \rightarrow U_\alpha \times E_{x_\alpha}$ by $\chi_\alpha(p) = (\pi(p), \varphi_{\alpha, \pi(p)}(p))$. Further we can choose arbitrary diffeomorphisms compatible with the G -structure of the bundle between each E_{x_α} and the typical fiber F . This allows us to treat χ_α 's as local trivializations. The transition functions between them are constant because they can be written in terms of these fixed diffeomorphisms and the transport diffeomorphisms $E_{x_\alpha} \rightarrow E_{x_\beta}$, which are entirely independent of any other points in the base. ■

An obvious consequence of the definition of parallel transport is that it endows the bundle with *unique horizontal path liftings*: given a path $\gamma : I \rightarrow M$ and a point $p \in \pi^{-1}(\gamma(0))$, its horizontal lifting is the smooth path $\tilde{\gamma}(t) = \text{tra}_{\gamma|_{[0,t]}}(p)$.

Definition 16.1.10 (Connection II: general connection). Given a G -bundle $(E \xrightarrow{\pi} M, G \curvearrowright F, \mathcal{G})$, a general connection on it is a smooth vector subbundle $\mathcal{H}E < TE$ such that $\mathcal{H}E \oplus \mathcal{V}E = TE$. The decomposition $T_p E = \mathcal{H}_p E \oplus \mathcal{V}_p E$ induces the natural projections

which we denote by two pairs of symbols:

$$T_p E \rightarrow \mathcal{H}_p E \oplus \mathcal{V}_p E, \quad X_p \mapsto (\text{hor } X_p, \text{ver } X_p). \quad (16.1.10)$$

For any smooth vector field $X \in \mathfrak{X}(E)$, the assignments $p \mapsto \text{hor } X_p$ and $p \mapsto \text{ver } X_p$ are also smooth vector fields, denoted $\text{hor } X$ and $\text{ver } X$, respectively.

The connection $\mathcal{H}E$ is called *flat* if it is integrable.

One thing to note immediately is that the map $X \mapsto \text{ver } X$ is $C^\infty(E)$ -linear and can therefore be identified with a $\mathcal{V}E$ -valued one-form on E . We denote it $\mathcal{V} \in \Omega^1(E; \mathcal{V}E)$. We have $\mathcal{H}E = \ker \mathcal{V}$. This leads to the third definition:

Definition 16.1.11 (Connection III: general connection). Given a **FB** $E \xrightarrow{\pi} M$, a general connection on it is a smooth one-form $\mathcal{V} \in \Omega^1(E; \mathcal{V}E)$ with values in the vertical bundle of E such that $\mathcal{V} \circ \mathcal{V} = \mathcal{V}$ and $\text{im } \mathcal{V} = \mathcal{V}E$ (so that it is a projection $TE \rightarrow \mathcal{V}E$). Given a general connection, we define the vertical and horizontal parts of a vector via

$$\text{ver } X = \mathcal{V}(X), \quad \text{hor } X = X - \mathcal{V}(X). \quad (16.1.11)$$

Exercise 16.1. Crucially, at this point we are not able to take exterior derivatives of connection forms, since they take values in a vector bundle $\mathcal{V}E$, and not a vector space. This situation will be remedied on principal bundles, whose vertical bundle is trivial and has a canonical trivialization (the Maurer-Cartan form).

Exercise 16.2. Show that definitions II and III are equivalent, i.e. there is a one-to-one correspondence between subbundles $\mathcal{H}E$ and one-forms \mathcal{V} .

Proposition 16.1.12. Let \mathcal{V} be a connection on $E \xrightarrow{\pi} M$. For any two horizontal vector fields $X, Y \in \Gamma(\mathcal{H}E)$, their “Lie bracket up to $\mathcal{H}E$ ”, i.e. the value of $[X, Y] \bmod H \in \Gamma(TE/\mathcal{H}E)$ at a point $p \in E$, is $C^\infty(E)$ -linear. In particular, the question of involutivity can be answered pointwise.

Proof. Let $\alpha \in \Omega^1(E)$. The chain rule and Cartan’s formula imply

$$i_{[X,Y]}\alpha = \mathcal{L}_X i_Y \alpha - i_Y \mathcal{L}_X \alpha = \mathcal{L}_X i_Y \alpha - i_Y(d(i_X \alpha)) - i_Y i_X d\alpha. \quad (16.1.12)$$

Suppose that for an open $U \subset E$ containing $p \in E$, we can find $\alpha_i \in \Omega^1(E|_U)$, $1 \leq i \leq l$, such that $\mathcal{H}E|_U = \bigcap_i \ker \alpha_i$. Then the fact that $X, Y \subset \mathcal{H}E$ implies that $i_X \alpha_i = 0$, so

$$\alpha_i([X, Y]) = -d\alpha_i(X, Y). \quad (16.1.13)$$

The value of the right hand side is $C^\infty(E)$ -linear in X and Y , and the left hand side uniquely determines $[X, Y] \bmod \mathcal{H}E$, since $\mathcal{H}E|_U = \bigcap_i \ker \alpha_i$.

What remains to show is that such α_i can indeed be chosen. Let $\pi' : \mathcal{V}E \rightarrow E$ denote the induced projection and let $\chi : \pi'^{-1}(U) \rightarrow U \times \mathbb{R}^k$ be a local trivialization. We let U shrink around p until it is contained in a chart $\kappa : U \rightarrow \mathbb{R}^n$. Then the components $\{\alpha_i\}_{i=1}^{n+k=l}$ of the map

$$TE \supset (\pi' \circ \mathcal{V})^{-1}(U) \xrightarrow{\alpha := (\kappa \times \text{id}_{\mathbb{R}^k}) \circ \chi \circ \mathcal{V}} \kappa(U) \times \mathbb{R}^k \subset \mathbb{R}^{n+k}, \quad (16.1.14)$$

where we view the connection form as a map $\mathcal{V} : TE \rightarrow \mathcal{V}E$, are precisely what we are looking for. ■

We have thus identified the following *tensor field* that quantifies the obstruction to the integrability of the horizontal distribution $\mathcal{H}E$.

Definition 16.1.13 (General curvature form). Given a general connection \mathcal{V} on a bundle E , the general curvature 2-form $\mathcal{R} \in \Omega^2(E; \mathcal{V}E)$ is defined by

$$\mathcal{R}(X, Y) := -\mathcal{V}([\text{hor } X, \text{hor } Y]) = -\mathcal{V}([X - \mathcal{V}(X), Y - \mathcal{V}(Y)]). \quad (16.1.15)$$

Its value is the unique representative of $[\text{hor } X, \text{hor } Y] \pmod{\mathcal{H}E}$ that lies in $\ker(\text{hor})$. The general connection \mathcal{V} is called *flat* if $\mathcal{R} = 0$ everywhere.

Exercise 16.3. Show that flatness (integrability) of $\mathcal{H}E$ is equivalent to flatness of \mathcal{V} and the existence of a flat general connection is equivalent to the flatness of the bundle E .

It follows from the definition of a connection that the restriction of π_{*p} to $\mathcal{H}_p E$ defines a linear isomorphism to $T_{\pi(p)} M$. This allows us to lift a vector $v \in T_{\pi(p)} M$ to a unique horizontal vector $v^h \in \mathcal{H}_p E$. Applying this pointwise to a vector field $X \in \mathfrak{X}(M)$, we construct the *horizontal lift* $X^h : E \rightarrow \mathcal{H}E \subset TE$. This is again a smooth vector field since it can be written as the composition

$$E \xrightarrow{\text{id}_E \times \pi_{TM}} E \times M \xrightarrow{\text{id}_E \times X} \pi^* TM \xrightarrow{i} \mathcal{H}E \subset TE, \quad (16.1.16)$$

where i is the inclusion, we understand the pullback bundle $\pi^* TM$ to be embedded in $E \times TM$, and where $\pi_{TM} : TM \rightarrow M$ is the bundle projection. How are the flows of X and X^h related? If γ is an integral curve of X^h , then $\pi \circ \gamma$ is clearly an integral curve of X , so we see that

$$\pi \circ F_t^{X^h} = F_t^X \circ \pi. \quad (16.1.17)$$

Therefore, for any two $X, Y \in \mathfrak{X}(M)$:

$$\pi_{*p}([X^h, Y^h](p)) = [X, Y](\pi(p)), \quad (16.1.18)$$

and we conclude

$$\mathcal{R}(X^h, Y^h) = -\mathcal{V}([X^h, Y^h]) = [X, Y]^h - [X^h, Y^h]. \quad (16.1.19)$$

We have proven the following.

Corollary 16.1.14. *Using the identification $\pi^*TM \xrightarrow{i} \mathcal{H}E \subset TE$, we can view \mathcal{R} as a map $\pi^*(\wedge^2 M) \rightarrow \mathcal{V}E$ given by*

$$(X_m, Y_m, p) \mapsto (X, Y, p) \mapsto [X^h, Y^h](p) - [X, Y]^h(p), \quad (16.1.20)$$

where $X, Y \in \mathfrak{X}(M)$ are arbitrary vector fields on M such that $X(m) = X_m$ and $Y(m) = Y_m$, and $m = \pi(p)$.

Let us now take a look at the local representatives of the connection form. Let (U_α, χ_α) be a local trivialization of our bundle $E \xrightarrow{\pi} M$ around a point $p \in E$, and let F be the typical fiber. Then in this trivialization

$$\begin{aligned} (\chi_\alpha)_*(\mathcal{H}E|_{U_\alpha}) &= (\chi_\alpha)_* \circ i(\pi^*(TM|_{U_\alpha})) = \{(\chi_\alpha)_* \circ i(X, p) \mid \pi_{TM}(X) = \pi(p)\} = \\ &= \{(X, -\Gamma^\alpha(X, f)) \in TU_\alpha \times TF \mid (X, f) \in TM|_{U_\alpha} \times F\} \end{aligned} \quad (16.1.21)$$

for some smooth map $\Gamma^\alpha : TM|_{U_\alpha} \times F \rightarrow TF$ known as the *Christoffel form* for this trivialization. More explicitly,

$$\Gamma^\alpha(X, f) = \text{pr}_2 \circ (\chi_\alpha)_* \circ \mathcal{V} \circ (\chi_\alpha)_*^{-1}(X, 0), \quad (16.1.22)$$

where $\text{pr}_2 : TU_\alpha \times TF \rightarrow TF$ is the projection onto the second component and $0 \in T_f F$. Then the projection $\mathcal{V} : TE \rightarrow \mathcal{V}E$ is locally represented as follows:

$$\hat{\mathcal{V}}_\alpha = (\chi_\alpha)_* \circ \mathcal{V} \circ (\chi_\alpha)_*^{-1}(X, \xi) = \xi + \Gamma^\alpha(X, \pi_{TF}(\xi)), \quad (X, \xi) \in TU_\alpha \times TF. \quad (16.1.23)$$

Finally it is essential to discuss how connections interact with pullbacks of bundles. Recall that categorical pullbacks of smooth fiber bundles exist due to the fact that a fiber bundle is a submersion, and a submersion is transversal to any smooth map (since the topological pullback of a pair of transversal maps $(f, \pi) : N \times E \rightarrow M$ is a smooth submanifold of $N \times E$).

Proposition 16.1.15 (Pullback connection [44, 1.2.7]). *Let $E \xrightarrow{\pi} M$ be a **FB** and $f : N \rightarrow M$ a smooth map. Let $\pi' : f^*E \rightarrow M$ be the pullback bundle and let $\tilde{f} : f^*E \rightarrow E$ be its other natural projection. Let E be equipped with a general connection \mathcal{V} . Then*

the pullback induces a unique connection denoted (confusingly) $f^*\mathcal{V}$ on f^*E such that $\tilde{f}_* \circ (f^*\mathcal{V}) = \mathcal{V} \circ \tilde{f}_*$.

Proof. First, why is $f^*\mathcal{V}$ a connection, i.e. a projection $T(f^*E) \rightarrow V(f^*E)$? The main observation is that $T_q\tilde{f}$ restricts to a linear isomorphism

$$V_q(f^*E) = T_q((f^*E)_{\pi'(q)}) \rightarrow T_{\tilde{f}(q)}(E_{\pi(\tilde{f}(q))}) = V_{\tilde{f}(q)}E. \quad (16.1.24)$$

Indeed, \tilde{f} is a diffeomorphism of fibers $(f^*E)_{\pi'(q)} \rightarrow E_{\pi(\tilde{f}(q))}$. This means that the following commuting square uniquely defines $f^*\mathcal{V}$:

$$(\mathcal{V})_{\tilde{f}(q)} \circ \tilde{f}_{*q} = \tilde{f}_{*q} \circ (f^*\mathcal{V})_q. \quad (16.1.25)$$

Finally, note that $\tilde{f}_{*q} \circ (f^*\mathcal{V})_q^2 = \mathcal{V}_{\tilde{f}(q)} \circ \tilde{f}_{*q} \circ (f^*\mathcal{V})_q = (\mathcal{V})_{\tilde{f}(q)}^2 \circ \tilde{f}_{*q} = (\mathcal{V})_{\tilde{f}(q)} \circ \tilde{f}_{*q} = \tilde{f}_{*q} \circ (f^*\mathcal{V})_q$, and therefore, since \tilde{f}_{*q} is an isomorphism, $(f^*\mathcal{V})^2 = f^*\mathcal{V}$. ■

Remark 16.1.16. Given two fiber bundles E_1, E_2 over M with fiber product $E_1 \xleftarrow{\pi^1} E_1 \times_M E_2 \xrightarrow{\pi^2} E_2$, and respective connections \mathcal{V}_1 and \mathcal{V}_2 on them, we can construct a connection $\mathcal{V}_1 \times_M \mathcal{V}_2 := \mathcal{V} = \pi^{1*}\mathcal{V}_1 + \pi^{2*}\mathcal{V}_2$ on $E_1 \times_M E_2$, called the *product connection*.

Finally, let us connect general connections to parallel transport, since that is the most intuitive (and, in a way, “correct”) way of defining a connection. We have seen that each tangent vector on M has a unique horizontal lift X^h under a given general connection. Therefore we can attempt to lift entire curves $\gamma : I \rightarrow M$ by lifting each velocity vector $\dot{\gamma}(t)$ individually and solving the resulting differential equation.

We now examine to what extent a connection \mathcal{V} determines parallel transports.

Theorem 16.1.17 (Parallel transport). *Let $\mathcal{H}E$ be a general connection on a fiber bundle $E \xrightarrow{\pi} M$ and let $\gamma : (a, b) \rightarrow M$ be a smooth path with $0 \in (a, b)$. Then there exists an open neighborhood $\mathcal{D}_\gamma \in E \times I$ of $E_{\gamma(0)} \times \{0\}$ and a smooth map $T_\gamma : \mathcal{D}_\gamma \rightarrow E$ such that for all $(p, t) \in \mathcal{D}_\gamma$:*

- (1) $\pi(T_\gamma(p, t)) = \gamma(t)$;
- (2) $\frac{d}{dt}T_\gamma(p, t) \in \mathcal{H}E$;
- (3) *T is reparametrization invariant: if $\phi : (a', b') \rightarrow (a, b)$ is smooth with $0 \in (a', b')$, then for all $s \in (a', b')$,*

$$T_\gamma(p, \phi(s)) = T_{\gamma \circ \phi}(T_\gamma(p, \phi(0)), s). \quad (16.1.26)$$

Proof. The standard proof is the following. Let F be the typical fiber and let $\chi_\alpha : \pi^{-1}(U_\alpha) \rightarrow U_\alpha \times F$ be a local trivialization for E around $\gamma(0)$. Property (2) can equivalently be stated as demanding that $\mathcal{V}(\frac{d}{dt}T_\gamma(p, t)) = 0$. In the local chart this condition becomes

$$\frac{d}{dt}\Phi_\gamma(p', t) = -\Gamma^\alpha(\dot{\gamma}(t), \Phi_\gamma(p', t)), \quad (16.1.27)$$

where $\Phi_\gamma(p', t) = \text{pr}_2 \circ \chi_\alpha \circ T_\gamma(\chi_\alpha^{-1}(p'), t) \in F$. Now, the right hand side of this equation is a smooth time-dependent vector field, therefore from general theory of ODEs this equation has a unique maximal integral solution $\Phi_\gamma : U'_\gamma \rightarrow U_\alpha \times F$ on some open domain $D'_\gamma \subset U_\alpha \times F \times I$. Since χ_α is a diffeomorphism, we obtain a map $T_\gamma : \mathcal{D}_\gamma \rightarrow E$ by $T_\gamma(p, t) = \chi_\alpha^{-1}(\gamma(t), \Phi_\gamma(\chi_\alpha(p), t))$.

Property (3) is also easily checked. Indeed, write $\gamma^h(t) := T_\gamma(p, t)$ for a horizontal lift of γ . Then $\frac{d}{ds}(\gamma^h \circ \phi(s)) = \dot{\phi}(s) \cdot \gamma^{h''}(\phi(s))$, and $\pi \circ \gamma^h \circ \phi = \gamma \circ \phi$, so $\gamma^h \circ \phi$ is a horizontal lift of $\gamma \circ \phi$. Finally, $\gamma^h \circ \phi(0) = T_\gamma(p, \phi(0))$, so the conclusion follows from uniqueness. ■

For this to extend to a proper parallel transport functor, we need to be able to lift paths globally. This leads us to the following definition.

Definition 16.1.18 (Connection IV: Ehresmann connection). A general connection $\mathcal{H}E$ on a bundle $E \xrightarrow{\pi} M$ is called an Ehresmann connection if it is *complete* in the sense that the parallel transport T_γ for any smooth curve $\gamma : (a, b) \rightarrow M$ (with $0 \in (a, b)$) is defined on all of $E_{\gamma(0)} \times (a, b)$.

We provide the following theorem without proof.

Theorem 16.1.19 ([45]). *Every FB admits a complete connection.*

Example 16.1.20. Every general connection on a bundle with compact fibers is complete.

Exercise 16.4. Show that definitions I and IV are equivalent, i.e. there is a one-to-one correspondence between Ehresmann connections and smooth parallel transport functors.

Note that parallel transport in fact only requires the path γ to be C^1 -differentiable. Let $C(m)$ be the group of piecewise smooth loops starting and ending at $m \in M$. The parallel transport along $\gamma \in C(m)$ yields an automorphism of the fiber E_m . For the trivial curve it coincides with identity. Assuming the parallel transport respects the structure group, the resulting map $C(m) \rightarrow G$ is called the *holonomy representation*. The image of this homomorphism forms a subgroup of the structure group. Further, let $C^0(m) \subset C(m)$ be the subgroup of nullhomotopic closed curves.

Definition 16.1.21 (Holonomy group). Given an Ehresmann connection $\mathcal{H}E$ on $E \xrightarrow{\pi} M$, the group of parallel transports along elements of $C(m)$ is called the holonomy group with basepoint m . It will be denoted $\text{Hol}_m(\mathcal{H}E)$. The restricted holonomy group is the subgroup $\text{Hol}_m^0(\mathcal{H}E) < \text{Hol}_m(\mathcal{H}E)$ generated by parallel transports along elements of $C^0(m)$.

Definition 16.1.22 (Monodromy of flat connections). In the special case that $\mathcal{H}E$ is flat (integrable), parallel transport is homotopy-invariant, and then we also get a prepresentation of the fundamental group $\pi_1(M, m) \rightarrow G$ called the *monodromy representation*. Its image is the *monodromy group*, denoted $\text{Mon}_m(\mathcal{H}E)$.

Regardless of whether a connection is complete, we can always expect parallel transport to be possible for loops contained in a sufficiently small neighborhood in the base, and a fixed initial condition $p \in E$. This allows us to describe curvature in terms of *infinitesimal holonomy*.

Proposition 16.1.23 (Curvature as infinitesimal holonomy). *Let $E \xrightarrow{\pi} M$ be a FB and $\mathcal{H}E$ a connection with curvature tensor \mathcal{R} . If $X, Y \in \mathfrak{X}(M)$ are two commuting vector fields, $[X, Y] = 0$, then*

$$R(X, Y)(p) = \frac{d^2}{dt ds} \left(F_{-t}^{X^h} \circ F_{-s}^{Y^h} \circ F_t^{X^h} \circ F_s^{Y^h}(p) \right) \Big|_{t=s=0} \quad (16.1.28)$$

Proof. Using Corollary 16.1.14, combined with $[X, Y] = 0$, we have $\mathcal{R}(X, Y)(p) = -[X^h, Y^h](p)$. The claim then follows from the well-known identity (9.4.11) for the Lie bracket of vector fields in terms of their flows. ■

Corollary 16.1.24. *A general connection is flat iff all of its restricted holonomy groups are trivial.*

16.2 Horizontal forms

Before we discuss connections on principal bundles, it will be useful to introduce the concept of horizontal differential forms. The following discussion is independent of any connections on bundles.

Definition 16.2.1 (Horizontal forms). Let $(P \rightarrow M, G \overset{\Phi}{\curvearrowright} P)$ be a (right) principal bundle and let $G \overset{\sigma}{\curvearrowright} V$ be a finite-dimensional representation of G . A V -valued differential k -form $\tilde{\eta} \in \Omega^k(P; V)$ is called horizontal of type σ if

1. $\tilde{\eta}|_{VP} = 0$, i.e. it is annihilated by any vertical vector (always true for $k = 0$);
2. $\Phi_g^* \tilde{\eta} = \sigma_{g^{-1}} \circ \tilde{\sigma}$ for every $g \in G$ (for a left principal bundle replace $\sigma_{g^{-1}}$ with σ_g).

The vector space of horizontal k -forms of type σ is denoted by $\Omega_{\text{hor}}^k(P; V)^\sigma$. Correspondingly, the space of ordinary horizontal k -forms is denoted $\Omega_{\text{hor}}^k(P)$.

Now consider the associated bundle $E := P^{[\sigma]} = P \times_G V$.

Proposition 16.2.2 ([17, Prop. 1.2.12]). *To every element $\tilde{\eta} \in \Omega_{\text{hor}}^k(P; V)^\sigma$ there corresponds a unique element $\eta \in \Omega^k(M; E)$ such that the following diagram commutes*

$$\begin{array}{ccc} \wedge^k P & \xrightarrow{\text{pr} \times \tilde{\eta}} & P \times V \\ \wedge^k \pi_* \downarrow & & \downarrow \iota \\ \wedge^k M & \xrightarrow{\eta} & E, \end{array} \quad (16.2.1)$$

where $\text{pr} : \wedge^k P \rightarrow P$ denotes the natural projection. The assignment $\tilde{\eta} \mapsto \eta$ defines a vector space isomorphism $\Omega_{\text{hor}}^k(P; V)^\sigma \rightarrow \Omega^k(M; E)$.

Proof. Let $m \in M$ and $X_i \in T_m M$, $i = 1, \dots, k$. Choose $p \in P$ fulfilling $\pi(p) = m$ and $Y_i \in T_p P$ such that $\pi_*(Y_i) = X_i$. We define

$$\eta_m(X_1, \dots, X_k) := \iota_p \circ \tilde{\eta}_p(Y_1, \dots, Y_k). \quad (16.2.2)$$

We must show that this definition does neither depend on the choice of p nor on the choice of the Y_i . Thus, take $p' = \Phi_g(p)$ and tangent vectors Y'_i at p' which also project onto the X_i . Then, there exist vertical vectors $Z_i \in T_{p'} P$ such that $Y'_i = \Phi_{g*}(Y_i) + Z_i$ and we obtain

$$\begin{aligned} \iota_{p'} \circ \tilde{\eta}_{p'}(Y'_1, \dots, Y'_k) &= \iota_{p \cdot g} \circ \tilde{\eta}_{p \cdot g}(\Phi_{g*}(Y_1) + Z_1, \dots, \Phi_{g*}(Y_k) + Z_k) = \\ &= \iota_p \circ \sigma_g \circ (\Phi_g^* \tilde{\eta})_p(Y_1, \dots, Y_k) = \iota_p \circ \tilde{\eta}_p(Y_1, \dots, Y_k). \end{aligned} \quad (16.2.3)$$

Here, we have used $\iota_{p \cdot g} = \iota_p \circ \sigma_g$ together with the horizontality and equivariance of $\tilde{\eta}$. Bijectivity and linearity of the assignment $\tilde{\eta} \mapsto \eta$ follows from the bijectivity and linearity of ι_p . ■

Note that, conversely, we have

$$\tilde{\eta}_p = \iota_p^{-1} \circ (\pi^* \eta)_p. \quad (16.2.4)$$

Remark 16.2.3. Note that elements $\tilde{\eta} \in \Omega_{\text{hor}}^\bullet(P; V)^\sigma$ and $\eta \in \Omega^\bullet(M; E)$ have a well-defined wedge product with any regular form $\beta \in \Omega^\bullet(M)$ on M , since $\tilde{\beta} \wedge \tilde{\eta} = \widetilde{\beta \wedge \eta}$. Thus both of these spaces are $\Omega^\bullet(M)$ -modules.

Let us also study local representations of horizontal forms. Let (U_α, χ_α) be a local trivialization of P , let $\tilde{s}_\alpha : P|_{U_\alpha} \rightarrow G$ be the corresponding equivariant map and let $s_\alpha : U \rightarrow P$ be the corresponding local section. We define the local representative of $\tilde{\eta} \in \Omega_{\text{hor}}^k(P; V)^\sigma$ by

$$\tilde{\eta}_\alpha := s_\alpha^* \tilde{\eta}. \quad (16.2.5)$$

This is a V -valued k -form on U . The following proposition shows that a horizontal form of type σ may be reconstructed from its local representatives.

Proposition 16.2.4 ([17, Prop. 1.2.14]). *Let $\tilde{\eta} \in \Omega_{\text{hor}}^k(P; V)$ and let $\tilde{\eta}_\alpha$ be its representative in a local trivialization (U_α, χ_α) given by (16.2.5). Then, for every $p \in P|_{U_\alpha}$, we have*

$$\tilde{\eta}_p = \sigma_{\tilde{s}_\alpha(p)^{-1}} \circ (\pi^* \tilde{\eta}_\alpha)_p. \quad (16.2.6)$$

Proof. By the equivariance of $\tilde{\eta}$, for every $p \in \pi^{-1}(U)$ and $Y_i \in T_p P$, we obtain

$$\begin{aligned} \sigma_{\tilde{s}_\alpha(p)} \circ \tilde{\eta}_p(Y_1, \dots, Y_k) &= \left(\Phi_{\tilde{s}_\alpha(p)^{-1}}^* \tilde{\eta} \right)_p(Y_1, \dots, Y_k) = \\ &= \tilde{\eta}_{p \cdot \tilde{s}_\alpha(p)^{-1}} \left(\Phi_{\tilde{s}_\alpha(p)^{-1}*}(Y_1), \dots, \Phi_{\tilde{s}_\alpha(p)^{-1}*}(Y_k) \right). \end{aligned} \quad (16.2.7)$$

Since $p \cdot \tilde{s}_\alpha(p)^{-1} = s_\alpha(\pi(p))$, we have $\Phi_{\tilde{s}_\alpha(p)^{-1}*}(Y_i) \in T_{s_\alpha(\pi(p))} P$ and

$$\pi_* \left(\Phi_{\tilde{s}_\alpha(p)^{-1}*}(Y_i) - s_{\alpha*} \circ \pi_*(Y_i) \right) = 0. \quad (16.2.8)$$

Thus, using the horizontality of $\tilde{\eta}$, in the above formula we may replace the tangent vectors $\Phi_{\tilde{s}_\alpha(p)^{-1}*}(Y_i)$ by $s_{\alpha*} \circ \pi_*(Y_i)$. This yields

$$\begin{aligned} \sigma_{\tilde{s}_\alpha(p)} \circ \tilde{\eta}_p(Y_1, \dots, Y_k) &= \tilde{\eta}_{s_\alpha(\pi(p))} (s_{\alpha*} \circ \pi_*(Y_1), \dots, s_{\alpha*} \circ \pi_*(Y_k)) = \\ &= (\pi^*(s_\alpha^* \tilde{\eta}))_p(Y_1, \dots, Y_k), \end{aligned} \quad (16.2.9)$$

and thus the assertion. ■

Remark 16.2.5. 1. If the transition functions on P are $t_{\beta\alpha}$ on $U_{\alpha\beta} = U_\alpha \cap U_\beta$, then

$$\tilde{\eta}_m^\beta = \sigma_{t_{\beta\alpha}(m)} (\tilde{\eta}_m^\alpha). \quad (16.2.10)$$

In other words, they satisfy the same transformation law as the sections of P itself. It is easy to show that any system of k -forms $\{\tilde{\eta}_\alpha\}$ fulfilling these conditions defines a unique element of $\Omega_{\text{hor}}^k(P; V)^\sigma$ with these local representatives.

2. If $\eta \in \Omega^k(M; E)$ corresponds to $\tilde{\eta}$, and (U_α, ξ_α) is the local trivialization of E induced

by (U_α, χ_α) via (14.2.11). We can define the local representative of η by

$$\eta_\alpha = \text{pr}_2 \circ \xi_\alpha \circ \eta|_U. \quad (16.2.11)$$

Using (16.2.2), (16.2.6) and (14.2.11), we compute

$$(\eta_\alpha)_m(X_1, \dots, X_k) = [(s_\alpha(m), (\tilde{\eta}_\alpha)_m(X_1, \dots, X_k))] = \xi_\alpha^{-1}(m, (\tilde{\eta}_\alpha)_m(X_1, \dots, X_k)) \quad (16.2.12)$$

for $m \in U$ and $X_i \in T_m M$. Thus, $\eta_\alpha = \tilde{\eta}_\alpha$ as expected.

3. In particular, for a section $\rho \in \Gamma^\infty(E) = \Omega^0(M; E)$, its representative $\rho^\alpha = \tilde{\rho}^\alpha \in C^\infty(U, V)$ is reconstructed via

$$\rho(m) = [(s(m), \rho^\alpha(m))], \quad \tilde{\rho}(p) = \sigma_{\tilde{s}_\alpha(p)^{-1}} \rho^\alpha(\pi(p)), \quad m \in U, p \in \pi^{-1}(U). \quad (16.2.13)$$

Example 16.2.6 (Tangent bundle of a Lie group, part 3). We pick up Example 14.2.19, where we considered the principal bundle $P = G \times G \rightarrow G$ given by $\pi(g_1, g_2) = g_1 g_2^{-1}$. We have

$$\pi_*(Y_1 + Y_2) = R_{g_2*}^{-1} Y_1 - L_{g_1 g_2^{-1}} R_{g_2*}^{-1} Y_2 = L_{g_1*} R_{g_2*}^{-1} \circ (\text{pr}_1^* \Theta_G - \text{pr}_2^* \Theta_G)(Y_1 + Y_2). \quad (16.2.14)$$

In Example 14.2.14 we have also defined the isomorphism $\vartheta : TG \rightarrow P \times_G \mathfrak{g}$, which can be viewed as a 1-form $\vartheta \in \Omega^1(G; P \times_G \mathfrak{g})$. By Proposition 16.2.2, it corresponds to a unique 1-form $\theta = \tilde{\vartheta} \in \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}$. We now compute it explicitly on a vector $Y \in T_{(g_1, g_2)} P$, which can be uniquely decomposed as $Y = Y_1 + Y_2$ with $(Y_1, Y_2) \in T_{g_1} G \times T_{g_2} G$:

$$\begin{aligned} \theta(Y) &= \iota_{(g_1, g_2)}^{-1} \circ (\pi^* \vartheta)(Y) = \iota_{(g_1, g_2)}^{-1} \circ \vartheta_{(g_1 g_2^{-1})} \left(L_{g_1*} R_{g_2*}^{-1} \left(L_{g_1*}^{-1} Y_1 - L_{g_2*}^{-1} Y_2 \right) \right) = \\ &= \iota_{(g_1, g_2)}^{-1} \left[(g_1 g_2^{-1}, e), \text{Ad}_{g_2} \left(L_{g_1*}^{-1} Y_1 - L_{g_2*}^{-1} Y_2 \right) \right] = \\ &= \iota_{(g_1, g_2)}^{-1} [(g_1, g_2), \Theta_G(Y_1) - \Theta_G(Y_2)] = \Theta_G(Y_1) - \Theta_G(Y_2). \end{aligned} \quad (16.2.15)$$

Therefore

$$\theta = \text{pr}_1^* \Theta_G - \text{pr}_2^* \Theta_G. \quad (16.2.16)$$

From the properties of the Maurer-Cartan form it is easy to see that this is indeed a horizontal form of type Ad. We will use it in Example 16.8.15.

16.3 Principal connections

Note that our definition of general connections so far hasn't taken into account any G -structure that might exist on the bundle. Only the parallel transport definition I of Ehresmann

connections was amenable to G -structures. It is not clear, however, how to adapt the definition of a general connection in terms of the horizontal bundle to a G -structure (other than indirectly via requiring parallel transports to act by morphisms of G -structures). To understand how to construct connections adapted to arbitrary G -structures, we will first study equivariant connections on principal bundles, and then learn to use them to induce connections on associated bundles.

We start by examining general connections on principal bundles before we narrow down a natural special class of such connections. In this section we assume that $(P \xrightarrow{\pi} M, G, \Phi)$ is a right principal G -bundle with the principal right action Φ . This action induces an action of the Lie algebra $\mathfrak{g} = \text{Lie } G$, given by the homomorphism $\mathfrak{g} \rightarrow \mathfrak{X}(P)$, $X \mapsto X_*$ (see Section 13.4). Recall that the space $\Gamma(VP)$ of vertical vector fields forms a Lie subalgebra of $\mathfrak{X}(P)$. Since the orbits of Φ are exactly the fibers of P , we have actually constructed a Lie algebra homomorphism $\mathfrak{g} \rightarrow \Gamma(VP)$. As we now show, this map actually trivializes VP , much in the same way as the Maurer-Cartan form trivializes the tangent bundle of a Lie group. The triviality of VP is a key property of principal bundles.

Proposition 16.3.1 ([17, Lem. 1.3.1]). *The map $\psi : P \times \mathfrak{g} \rightarrow VP$ given by $(p, A) \mapsto A_*(p)$ is an isomorphism of VBs over P . Hence, ψ^{-1} is a global trivialization for VP .*

Proof. ψ clearly respects the fibers since $A_*(p) \in \mathcal{V}_p P$, so it covers the identity on P . It is also fiberwise linear. Now let $p \in P$ and (U, χ) be a trivialization of P around p . We have $\psi(p, A) = \Phi_*^p(A)$. Note that Φ^p is a diffeomorphism $G \rightarrow P_p$. Therefore $\Phi_*^p : \mathfrak{g} \rightarrow \mathcal{V}_p P$ is an isomorphism of vector spaces. ψ is obviously smooth, as $\psi(p, A) = \Phi_*^p(A) = \Phi_*(0_p, A)$, where $0_p \in T_p P$ is the origin. Since ψ is a smooth vertical bundle morphism that is a fiberwise diffeomorphism, it is a bundle isomorphism. ■

This allows us to interpret the connection 1-form $\mathcal{V} \in \Omega^1(P; VP)$ and the curvature 2-form $\mathcal{R} \in \Omega^2(P; VP)$ as \mathfrak{g} -valued forms

$$\omega := \psi^{-1} \circ \mathcal{V} \in \Omega^1(P; \mathfrak{g}), \quad \Omega^\omega := \psi^{-1} \circ \mathcal{R} \in \Omega^2(P; \mathfrak{g}). \quad (16.3.1)$$

We also note for the future that for horizontal vectors we have

$$\text{ver}([X^h, Y^h])_p = -\Phi_*^p(\Omega^\omega(X^h, Y^h)). \quad (16.3.2)$$

Now we find the necessary and sufficient condition for such a 1-form ω to correspond to a general connection.

Proposition 16.3.2. *A 1-form $\omega \in \Omega^1(P; \mathfrak{g})$ defines a general connection on P iff $\omega(A_*) = A$ for any $A \in \mathfrak{g}$.*

Proof. Let $A \in \mathfrak{g}$ and suppose $\omega(A_*)_* = \mathcal{V}(A_*)$ (all asterisks here denote taking the corresponding Killing vector field on P). Then $\omega(A_*)_* = \mathcal{V}(A_*) = A_*$. The same claim says that infinitesimal action corresponding to the principal action Φ defines a linear isomorphism $\mathfrak{g} \rightarrow \mathcal{V}_p P$ for each $p \in P$, so we conclude that $\omega(A_*) = A$.

Conversely, note that the post-composition with ψ shows that $\Omega^1(P; \mathfrak{g})$ is in one-to-one correspondence with bundle morphisms $TP \rightarrow VP$ over P . Saying that $\omega \in \Omega^1(P; \mathfrak{g})$ then precisely comes down to demanding that $\mathcal{V} := \psi \circ \omega$ is idempotent ($\mathcal{V} \circ \mathcal{V} = \mathcal{V}$). Now, $\omega(A_*) = A$ means that $\omega \circ \psi = \text{id}_{\mathfrak{g}}$, so the claim follows. ■

Corollary 16.3.3. $\Omega^\omega(X, Y) = -\omega([\text{hor } X, \text{hor } Y])$.

Proof. Indeed, $-\omega([\text{hor } X, \text{hor } Y])_* = -\mathcal{V}([\text{hor } X, \text{hor } Y]) = \mathcal{R}(X, Y) = \Omega^\omega(X, Y)_*$. In particular, Ω^ω is horizontal. ■

Obviously we are interested not in general connections on P , but in *G-equivariant* ones – we want the distribution $\mathcal{H}P$ to be invariant under the action Φ :

$$\Phi_{g*}(\mathcal{H}_p P) = \mathcal{H}_{\Phi_g(p)} P, \quad \mathcal{V} \circ \Phi_{g*} = \Phi_{g*} \circ \mathcal{V}. \quad (16.3.3)$$

Let us understand this condition in terms of the 1-form ω .

Proposition 16.3.4. *Let $\omega \in \Omega^1(P; \mathfrak{g})$ be such that $\omega(X_*) = X$ for all $X \in \mathfrak{g}$, and let \mathcal{V} be the corresponding general connection on P . Then \mathcal{V} is G-equivariant iff ω is adjoint-equivariant in the sense that*

$$\Phi_g^* \omega = \text{Ad}_g^{-1} \circ \omega. \quad (16.3.4)$$

(on a left principal bundle replace Ad_g^{-1} with Ad_g).

Proof. We have

$$\Phi_g^* \omega = \Phi_g^* \circ \psi^{-1} \circ \mathcal{V} = \psi^{-1} \circ \mathcal{V} \circ \Phi_{g*} = \psi^{-1} \circ \Phi_{g*} \circ \mathcal{V} = \psi^{-1} \Phi_{g*} \psi \circ \omega, \quad (16.3.5)$$

and according to Proposition 13.4.2, for a right action Φ ,

$$\psi^{-1} \Phi_{g*} \psi(X) = \text{Ad}_g^{-1} X. \quad (16.3.6)$$

■

We can thus give the following definition.

Definition 16.3.5 (Principal connection). A principal connection on a (right) **PFB** ($P \xrightarrow{\pi} M, G, \Phi$) is a general connection $\mathcal{H}P$ that is equivariant in the following sense:

$$\Phi_{g*}(\mathcal{H}_p P) = \mathcal{H}_{\Phi_g(p)} P. \quad (16.3.7)$$

A *principal connection form* on P is a \mathfrak{g} -valued 1-form $\omega \in \Omega^1(P; \mathfrak{g})$ such that

1. $\omega(A_*) = A$ for all $A \in \mathfrak{g}$.
2. $\Phi_g^*\omega = \text{Ad}_g^{-1} \circ \omega$ for all $g \in G$ (replace Ad_g^{-1} with Ad_g on a left **PFB**).

There is a one-to-one correspondence between principal connections and principal connection forms given by

$$\mathcal{H}P = \ker \omega. \quad (16.3.8)$$

The space of all principal connection forms is denoted $\mathcal{A}(P)$ (with topology induced from $\Omega^1(P; \mathfrak{g})$).

Proposition 16.3.6. *Every **PFB** admits a principal connection.*

Proof. Choose a countable, locally finite system of local trivializations $\{(U_i, \chi_i)\}$ of P and a subordinate **POU** $\{f_i\}$ on M . At every point $p = \chi_i^{-1}(m, e)$, $m \in M$, define

$$\mathcal{H}_p P := \chi_{i*}^{-1}(\mathcal{T}_{(m,e)}(U_i \times \{e\})). \quad (16.3.9)$$

Clearly, $\mathcal{H}_p P$ is complementary to $\mathcal{V}_p P$. If we transport this subspace by Φ_{g*} , $g \in G$, to the remaining points of the fiber P_m , for every $m \in U_i$, then we obtain a principal connection on the trivial principal bundle $\pi^{-1}(U_i)$. Denote the corresponding connection form by $\tilde{\omega}_i$. We defin the following family of \mathfrak{g} -valued 1-forms on P :

$$(\omega_i)_p := \begin{cases} 0, & p \notin \pi^{-1}(U_i), \\ (\pi^* f_i) \tilde{\omega}_i, & p \in \pi^{-1}(U_i). \end{cases} \quad (16.3.10)$$

Since the collection of sets $\{\text{supp } f_i\}$ is locally finite, $\omega := \sum_i \omega_i$ is a well-defined smooth \mathfrak{g} -valued 1-form on P . It remains to check the two properties necessary for it to be a principal connection form.

For $p \in P$ and $A \in \mathfrak{g}$, we have

$$\omega_p(A_*(p)) = \sum_{i \in I_p} (\omega_i)_p(A_*(p)) = \sum_{i \in I_p} (\omega_i)_p(A_*(p)), \quad (16.3.11)$$

where I_p is the set of indices i for which $\pi(p) \in U_i$. Since every $\tilde{\omega}_i$ is a connection form, for $i \in I_p$, we obtain

$$(\omega_i)_p(A_*(p)) = f_i(\pi(p))(\tilde{\omega}_i)_p(A_*(p)) = f_i(\pi(p))A. \quad (16.3.12)$$

Now, $\sum_{i \in I_p} f_i(\pi(p)) = 1$ implies that $\omega_p(A_*(p)) = A$.

For adjoint-equivariance, it suffices to verify it for every ω_i restricted to $\pi^{-1}(U_i)$. Since $\Phi_g^*((\pi^* f_i) \tilde{\omega}_i) = (\pi^* f_i)(\Phi_g^* \tilde{\omega}_i)$ and all $\tilde{\omega}_i$ are equivariant, the assertion follows. ■

Proposition 16.3.7. *Principal connections are complete.*

Proof. Let ω be a principal connection form on P and let $\gamma : I \rightarrow M$ be a C^1 -path. We want to show that parallel transport is defined globally, so, in the notation of 16.1.17, $\mathcal{D}_\gamma = I \times P$. Using local trivializations, it is easy to construct a path $\delta : I \rightarrow P$ such that $\pi \circ \delta = \gamma$ and $\delta'(0) = p$, given $p \in P_{\gamma(0)}$. We look for a smooth curve $g : I \rightarrow G$ such that $\gamma^h(t) := \delta(t) \cdot g(t)$ is a horizontal lift of γ : $\pi \circ \gamma^h = \gamma$ and $\dot{\gamma}^h(t) \in \mathcal{H}P$ for all $t \in I$.

We note that the principal right action preserves the fibers of P , so the first demand, that $\pi \circ \gamma^h = \gamma$, is immediately satisfied for any choice of $g(t)$. In terms of ω , the second condition on the lift reads

$$0 = \omega \left(\frac{d}{dt} (\delta(t) \cdot g(t)) \right). \quad (16.3.13)$$

Now, by the Leibniz rule, $\frac{d}{dt} (\delta(t) \cdot g(t)) = \Phi_{g(t)*}(\dot{\delta}(t)) + \Phi_*^{\delta(t)}(\dot{g}(t))$. By equivariance of ω , $\omega(\Phi_{g(t)*}(\dot{\delta}(t))) = \text{Ad}_{g(t)}^{-1}\omega(\dot{\delta}(t))$. Moreover, by the first defining property of ω , $\omega(\Phi_*^{\delta(t)}(\dot{g}(t))) = \omega((\dot{g}(t))_*(\delta(t))) = L_{g(t)*}^{-1}\dot{g}(t)$. Therefore the condition (16.3.13) reduces to

$$\dot{g}(t) = -R_{g(t)*}\omega(\dot{\delta}(t)). \quad (16.3.14)$$

We claim that this equation has a unique global solution $g : I \rightarrow G$, which is the integral curve of a time-dependent right-invariant vector field $X : G \times I \rightarrow T(G \times I) \cong TG \times I \times \mathbb{R}$ given by

$$(h, s) \mapsto X_s(h) := (-R_{h*}\omega(\dot{\delta}(s)), (\partial_t)_s) \in T_h G \times T_g I \cong T_h G \times \mathbb{R}. \quad (16.3.15)$$

Here, ∂_t denotes the canonical unit vector field in $T_s I \cong \mathbb{R}$. Moreover, right-invariance is immediate: $X_s(ha) = -R_{ha*}\omega(\dot{\delta}(s)) = R_{a*}(-R_{h*}\omega(\dot{\delta}(s))) = R_{a*}X_s(h)$. Obviously, solutions of (16.3.14) correspond precisely with integral curves of X . The assertion now follows from the fact that invariant vector fields on Lie groups are complete (Propositon 9.2.2). ■

Thus, principal connections are Ehresmann connections and define global parallel transport, which we denote by tra^ω . We will study it in more detail in Section 16.5.

Corollary 16.3.8. *If ω is a principal connection form, then $\Omega^\omega \in \Omega_{\text{hor}}^2(P; \mathfrak{g})^{\text{Ad}}$.*

Proof. From Corollary 16.3.3 it is obvious that Ω^ω is horizontal since $\text{hor } X = \text{hor } Y = 0$ if X, Y are vertical. To verify equivariance, recall that $\text{hor } X = X - \mathcal{V}(X)$ and note that $\mathcal{V} \circ \Phi_{g^*} = \Phi_{g^*} \circ \mathcal{V}$. Thus

$$(\Phi_g^* \Omega^\omega)(X, Y) = \Omega^\omega(\Phi_{g^*} X, \Phi_{g^*} Y) = -\omega(\Phi_{g^*}[\text{hor } X, \text{hor } Y]) = -(\Phi_g^* \omega)([\text{hor } X, \text{hor } Y]), \quad (16.3.16)$$

so equivariance follows from that of ω . ■

Remark 16.3.9. Any principal connection form $\omega \in \Omega^1(P; \mathfrak{g})$ is clearly not horizontal because it assigns value A to the vertical Killing vector field A_* . However, if ω and ω' are two arbitrary principal connection forms on P , then their difference is horizontal of type Ad:

$$(\omega - \omega') \in \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}. \quad (16.3.17)$$

Conversely, ω' is a principal connection form iff this is true. The crucial implication of this is that, while there is no natural vector space structure on the set $\mathcal{A}(P)$ of principal connections on P , it can be identified with an affine space modeled on $\Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}$ (recall that an affine space is nothing but a torsor of a vector space as an abelian group).

Lemma 16.3.10 ([17, Lem. 1.4.2]). *Let $P(M, G)$ be a principal bundle with a connection $\mathcal{H}P$, let A_* be a Killing vector field on P , let $X \in \mathfrak{X}(P)$ be horizontal and let $Y \in \mathfrak{X}(M)$. Then $[A_*, X]$ is horizontal and $[A_*, Y^h] = 0$.*

Proof. For any $p \in P$, we have

$$[A_*, X]_p = (\mathcal{L}_{A_*} X)_p = \frac{d}{dt} \Big|_0 ((\Phi_{e^{-tA}})_* X)_p. \quad (16.3.18)$$

Since X is horizontal, $(\Phi_{e^{-tA}})_* X$ is horizontal for all t . Thus $[A_*, X]$ is horizontal, too. To prove the second statement, recall that the horizontal lift Y^h is G -invariant, that is, the path $t \mapsto ((\Phi_{e^{-tA}})_* Y^h)_p$ is constant and equal to Y^h_p . This yields the assertion. ■

Proposition 16.3.11 (Structure Equation [17, Prop. 1.4.9]). *Let P be a principal bundle and ω a principal connection form on it with curvature form Ω^ω . Then*

$$\Omega^\omega = d\omega + \frac{1}{2} [\omega, \omega]. \quad (16.3.19)$$

Proof. We evaluate both sides on vector fields $X, Y \in \mathfrak{X}(P)$. By the definition of the bracket, we have $\frac{1}{2} [\omega, \omega](X, Y) = [\omega(X), \omega(Y)]$. Clearly, it is enough to consider the following three cases:

1. X and Y are horizontal. Then, $\omega(X) = \omega(Y) = 0$ and

$$d\omega(X, Y) = d\omega(\text{hor } X, \text{hor } Y) = d\omega(X, Y). \quad (16.3.20)$$

2. X is vertical and Y is horizontal. Then $\omega(Y) = 0$ and $\Omega(X, \cdot) = 0$, so the right hand side of (16.3.19) vanishes. To calculate the left hand side, without loss of generality, we may assume $X = A_*$ for some $A \in \mathfrak{g}$. Then, $\omega(A_*) = A$ and we obtain

$$d\omega(A_*, Y) = Y(\omega(A_*)) - A_*(\omega(Y)) - \omega([A_*, Y]) = -\omega([A_*, Y]) = 0, \quad (16.3.21)$$

because, according to Lemma 16.3.10, $[A_*, Y]$ is horizontal.

3. X and Y are vertical. Then $\Omega(X, Y) = 0$. Taking $X = A_*$ and $Y = B_*$ for some $A, B \in \mathfrak{g}$ and using $[A_*, B_*] = [A, B]_*$ (cf. Proposition 13.4.2), we calculate

$$d\omega(A_*, B_*) = -\omega([A_*, B_*]) = -\omega([A, B]_*) = -[A, B] = -[\omega(A_*), \omega(B_*)]. \quad (16.3.22)$$

■

Now let us look at local representations of connections. Let $s_\alpha \in \Gamma^\infty(P|_{U_\alpha})$ be a local section defining a local trivialization of P . The local representatives of ω and Ω^ω are defined by

$$\mathcal{A}_\alpha := s_\alpha^* \omega, \quad \mathcal{F}_\alpha := s_\alpha^* \Omega^\omega. \quad (16.3.23)$$

In gauge theories of physics, \mathcal{F}_α is precisely the tensor often called the *field strength*.

Remark 16.3.12. If (U, φ) is a local chart on M defining local coordinates $\{x^\mu\}_{\mu=1}^{\dim M}$ and $\{\mathbf{t}_a\}$ is a basis in \mathfrak{g} , then a local frame in the bundle of \mathfrak{g} -valued k -forms on M is given by

$$\{dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k} \otimes \mathbf{t}_a\}. \quad (16.3.24)$$

With respect to this frame, the local representatives \mathcal{A} and \mathcal{F} take the form (using Einstein notation)

$$\mathcal{A} = \mathcal{A}_\mu^a dx^\mu \otimes \mathbf{t}_a, \quad \mathcal{F} = \frac{1}{2} \mathcal{F}_{\mu\nu}^a dx^\mu \wedge dx^\nu \otimes \mathbf{t}_a. \quad (16.3.25)$$

The Structure Equation in local form reads

$$\mathcal{F} = d\mathcal{A} + \frac{1}{2} [\mathcal{A}, \mathcal{A}], \quad (16.3.26)$$

and in components

$$\mathcal{F}_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + c_{bc}^a \mathcal{A}_\mu^b \mathcal{A}_\nu^c, \quad (16.3.27)$$

where c_{bc}^a are the structure constants of \mathfrak{g} with respect to the basis $\{\mathbf{t}_a\}$.

In physics, local gauge potentials \mathcal{A} are usually defined in different units, and also so as to be Hermitian as opposed to skew-Hermitian in case of unitary structure groups, so that for example in electromagnetism \mathcal{A} is replaced by $\frac{ie}{\hbar c}$, where e, \hbar, c are the electron charge, Planck constant, and speed of light, respectively. More generally the physicist's connection form is $ig\mathcal{A}$, where $g \in \mathbb{R}$ is a *coupling constant*.

We now show how ω may be reconstructed from \mathcal{A}_α locally. For that purpose, recall that a section s_α defines an equivariant mapping $\tilde{s}_\alpha : P \rightarrow G$ by

$$(s_\alpha \circ \pi(p)) \cdot \tilde{s}_\alpha(p) = p, \quad p \in P. \quad (16.3.28)$$

Proposition 16.3.13 ([17, Prop. 1.3.11]). *In the above setting,*

$$\omega_p = \text{Ad}_{\tilde{s}_\alpha(p)}^{-1}(\pi^*\mathcal{A}_\alpha)_p + (\tilde{s}_\alpha^*\Theta_G)_p, \quad \Omega_p^\omega = \text{Ad}_{\tilde{s}_\alpha(p)}^{-1}(\pi^*\mathcal{F}_\alpha)_p, \quad (16.3.29)$$

with Θ_G denoting the Maurer-Cartan form on G .

Proof. We will omit the index α in the proof. Let $X \in T_p P$ and let $t \mapsto \gamma(t)$ be a curve representing X . Then, using (16.3.28), we compute

$$\begin{aligned} X = \dot{\gamma}(0) &= \left. \frac{d}{dt} \right|_0 ((s \circ \pi(\gamma(t))) \cdot \tilde{s}(\gamma(t))) = \\ &= (\Phi_{\tilde{s}(p)})_{*s(\pi(p))} \circ (s \circ \pi)_{*p}(X) + (\Phi_{s(\pi(p))})_{*\tilde{s}(p)} \circ \tilde{s}_{*p}(X). \end{aligned} \quad (16.3.30)$$

Denoting the first and the second summand by X^s and X^v , respectively, we get a decomposition $X = X^s + X^v$, where X^s is tangent to the submanifold $\Phi_{\tilde{s}(p)}(s(U))$ and X^v is vertical. We calculate

$$\begin{aligned} \omega_p(X^s) &= (\Phi_{\tilde{s}(p)}^*\omega)_{s(\pi(p))}((s \circ \pi)_{*p}(X)) = \\ &= \text{Ad}_{\tilde{s}(p)}^{-1}(s^*\omega)_{\pi(p)}(\pi_*(X)) = \text{Ad}_{\tilde{s}(p)}^{-1}(\pi^*\mathcal{A})_p(X). \end{aligned} \quad (16.3.31)$$

This yields the first summand in (16.3.29). On the other hand, by the definition of ω ,

$$\omega_p(X^v) = (\Phi_*^P)^{-1} \circ \left(\Phi^{s(\pi(p))} \right)_{\tilde{s}(p)} \circ \tilde{s}_{*p}(X). \quad (16.3.32)$$

Using the obvious identity $(\Phi^P)^{-1} \circ \Phi^{s(\pi(p))} \circ L_{\tilde{s}(p)} = \text{id}_G$, together with

$$(\tilde{s}^*\Theta_G)_p(X) = L_{\tilde{s}(p)^{-1}*} \circ \tilde{s}_{*p}(X), \quad (16.3.33)$$

we obtain $\omega_p(X) = (\tilde{s}^*\Theta_G)_p(X)$. ■

The following is immediate from $s_\alpha \cdot t_{\alpha\beta} = s_\beta$ or $\tilde{s}_\alpha = t_{\alpha\beta}\tilde{s}_\beta$. The result can be contrasted with the transformation law (16.2.10) for horizontal forms such as curvature.

Corollary 16.3.14 ([17, Cor. 1.3.12]). *If the transition functions on P are $\{t_{\beta\alpha}\}$, then the local representatives of ω are related as follows:*

$$(\mathcal{A}_\beta)_m = \text{Ad}_{t_{\alpha\beta}(m)}^{-1} \circ (\mathcal{A}_\alpha)_m + (t_{\alpha\beta}(m)^*)^* \Theta_G|_m, \quad (\mathcal{F}_\beta)_m = \text{Ad}_{t_{\alpha\beta}(m)}^{-1} \circ (\mathcal{F}_\alpha)_m. \quad (16.3.34)$$

Conversely, any system of Lie algebra-valued 1-forms $\{\mathcal{A}_\alpha\}$ fulfilling this transformation law defines a unique principal connection form ω with these local representatives.

Remark 16.3.15. In physics, this transformation law is often written as

$$\mathcal{A} \rightarrow g^{-1} \mathcal{A} g + g^{-1} dg, \quad (16.3.35)$$

where $g(m)$ is the matrix-valued transition function (*passive local gauge transformation*). Meanwhile local curvature transforms by simple conjugation. Notably, if the structure group is abelian, then all local representatives of Ω^ω agree with each other, thus defining a *global* curvature form $\mathbb{F} \in \Omega^1(M; \mathfrak{g})$ on M . This is why, for example, in electrostatics, which is a U_1 -gauge theory, the field strength \mathcal{F} ($\frac{ie}{\hbar c}$ times the magnetic field 2-form) is defined globally.

Example 16.3.16 (Canonical connection on the product bundle). Consider the trivial product bundle $P = M \times G \xrightarrow{\text{pr}_M} M$. Take the flat principal connection $\mathcal{H}P$ defined as

$$\mathcal{H}_{(m,g)} P := T_{(m,g)}(M \times \{g\}). \quad (16.3.36)$$

The connection form corresponding to $\mathcal{H}P$ is

$$\omega = \text{pr}_G^* \Theta_G, \quad (16.3.37)$$

where $\text{pr}_G : M \times G \rightarrow G$ denotes the canonical projection and Θ_G is the Maurer-Cartan form. The curvature vanishes, as expected, by the Structure Equation combined with the Maurer-Cartan equation.

Example 16.3.17 (Reductive Klein geometry). Let (G, H) be a Klein geometry (or, more broadly, a homogeneous space). Note that the projection $G \rightarrow G/H$ as a principal H -bundle doesn't carry a natural connection form. The Maurer-Cartan form doesn't satisfy the definition because it is not \mathfrak{h} -valued, where $\mathfrak{h} = \text{Lie } H$. To find a connection, assume, additionally, that (G, H) is *reductive*, that is, the Lie algebra of G admits a decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m} \quad (16.3.38)$$

as $\text{Ad}(H)$ -modules, meaning that $\text{Ad}(H)\mathfrak{m} \subset \mathfrak{m}$. (If G is semisimple, then \mathfrak{m} can be chosen to be the orthogonal complement to \mathfrak{h} in the sense of the Killing form.) In other words, \mathfrak{g} is a semidirect product.

Clearly, the vertical subspace at $g \in G$ is given by $L_{g*}(\mathfrak{h})$. Since for any $g \in G$, we have $T_g G = L_{g*}(\mathfrak{h}) \oplus L_{g*}(\mathfrak{m})$, the left invariant distribution $g \mapsto H_g G := L_{g*}(\mathfrak{m})$ on G is complementary to the canonical vertical bundle. Using the reductivity, it is easy to show that HG is H -invariant. Thus, HG defines a connection on G . The corresponding connection form is given by

$$\omega^0 = \text{pr}_{\mathfrak{h}} \circ \Theta_G, \quad (16.3.39)$$

where $\text{pr}_{\mathfrak{h}}$ is the projection $\mathfrak{g} \rightarrow \mathfrak{h}$ with respect to the above decomposition. Therefore the Maurer-Cartan form is decomposed into the sum

$$\Theta_G = \omega^0 + \theta^0, \quad (16.3.40)$$

where $\theta^0 = \text{pr}_{\mathfrak{m}} \circ \Theta_G$. We can now expand the Maurer-Cartan equation for Θ_G :

$$0 = d\Theta_G + \frac{1}{2} [\Theta_G, \theta_G] = d\omega^0 + \frac{1}{2} [\omega^0, \omega^0] + d\theta^0 + \frac{1}{2} [\theta^0, \theta^0] + [\omega^0, \theta^0]. \quad (16.3.41)$$

Since θ^0 is \mathfrak{m} -valued, by projecting onto \mathfrak{h} we get the following identity for the curvature $\Omega^0 = d\omega^0 + \frac{1}{2} [\omega^0, \omega^0]$:

$$\Omega^0 = -\frac{1}{2} \text{pr}_{\mathfrak{h}} [\theta^0, \theta^0] - \text{pr}_{\mathfrak{h}} [\omega^0, \theta^0]. \quad (16.3.42)$$

In general, this is not zero. However, in the special case that \mathfrak{m} is an ideal (and thus also a Lie subalgebra), both terms vanish and thus ω^0 is flat. As an example, the canonical connection on the affine space $A(\mathbb{R}^n) = \text{Aff}_n(\mathbb{R})/\text{GL}_n(\mathbb{R})$ is flat because in the reductive decomposition of the affine Lie algebra $\mathfrak{a}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R}) \oplus \mathbb{R}^n$, the second factor is an ideal since the translation subgroup is a normal subgroup of $\text{Aff}_n(\mathbb{R})$.

Example 16.3.18 (Canonical connection on the Stiefel bundle). Recall the Stiefel bundles

$$V_k(\mathbb{K}^n) = O_n/O_{n-k} \rightarrow G_k(\mathbb{K}^n) = O_n/(O_{n-k} \times O_k). \quad (16.3.43)$$

Denote the Lie algebra of the group O_i by \mathfrak{u}_i , $i = k, n-k, n$ (it is actually \mathfrak{so}_i , \mathfrak{u}_i , and \mathfrak{sp}_i in the real, complex, and quaternionic cases, respectively). Since O_{n-k} and O_k acts in complementary orthogonal subspace of \mathbb{K}^n , the direct sum of the Lie algebras is a Lie subalgebra of \mathfrak{u}_n and we have a direct sum decomposition

$$\mathfrak{u}_n = \mathfrak{u}_k \oplus \mathfrak{m}, \quad (16.3.44)$$

where $\mathfrak{m} = \mathfrak{u}_{n-k} \oplus \mathfrak{n}$ and \mathfrak{n} is the orthogonal complement of

$$\mathfrak{u}_k \oplus \mathfrak{u}_{n-k} \subset \mathfrak{u}_n \quad (16.3.45)$$

with respect to the Killing form. The restriction of the adjoint representation $\text{Ad}(\mathcal{O}_k)$ acts on \mathfrak{u}_{n-k} trivially and leaves the subspace \mathfrak{n} invariant. Thus, the above decomposition is reductive. As in Example 16.3.17, we put

$$\omega^c := \text{pr}_{\mathfrak{u}_k} \circ \Theta_G. \quad (16.3.46)$$

Clearly ω^c is a \mathfrak{u}_k -valued 1-form on \mathcal{O}_n . Since ω^c is invariant under the \mathcal{O}_{n-k} -action on \mathcal{O}_n , it descends to a \mathfrak{u}_k -valued 1-form on $V_k(\mathbb{K}^n)$ which we denote by the same symbol. We claim that ω^c is a connection form. To prove this, we have to check the defining properties. To check property 1, note that the Killing vector field of the right \mathcal{O}_k -action on \mathcal{O}_n generated by $A \in \mathfrak{u}_k$ coincides with A viewed as a left-invariant vector field,

$$(A_*)g = \frac{d}{dt} \Big|_0 (ge^{tA}) = gA, \quad g \in \mathcal{O}_n. \quad (16.3.47)$$

Since the right actions of \mathcal{O}_k and \mathcal{O}_{n-k} on \mathcal{O}_n commute, the Killing vector field of the right \mathcal{O}_k -action on $V_k(\mathbb{K}^n)$ generated by $A \in \mathfrak{u}_k$ may be identified with A_* . Now, condition 1 follows from the defining equation of the Maurer-Cartan form, $\Theta_G(A_*) = A$. Condition 2 follows immediately from the right \mathcal{O}_n -equivariance of Θ_G . The connection defined by ω^c is called the *canonical*, or *universal*, connection of the Stiefel bundle. By the left-invariance of Θ_G , the canonical connection is invariant under left translations of \mathcal{O}_n .

We give an explicit description of ω^c in terms of matrix-valued functions: let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis in \mathbb{K}^n . If we choose the k -frame $u_0 = (\mathbf{e}_1, \dots, \mathbf{e}_k)$, then the subgroups \mathcal{O}_k and \mathcal{O}_{n-k} are given in block matrix form by an upper diagonal ($k \times k$)-block and by a lower diagonal ($(n-k) \times (n-k)$)-block in \mathcal{O}_n , respectively. Let $g \in \mathcal{O}_n$ and let a_j^i be the corresponding $(n \times n)$ -matrix with respect to the standard basis. Since $a^\dagger a = I$, ω^c is represented by a $(k \times k)$ -valued 1-form on $V_k(\mathbb{K}^n)$,

$$(\omega^c)_\beta^\alpha = (a^\dagger)_j^\alpha d g^j_\beta, \quad (16.3.48)$$

where $\alpha, \beta = 1, \dots, k$ and $j = 1, \dots, n$. Denoting by u the matrix-valued function which assigns to the k -frame $u_\alpha = g_\alpha^j \mathbf{e}_j$ the $(n \times k)$ -matrix g_α^j , we obtain

$$\omega^c = u^\dagger du. \quad (16.3.49)$$

Since $g^\dagger g = I$, we have $u^\dagger u = I$.

Remark 16.3.19. In the above realization, the horizontal vectors of ω^c at the point $p_0 = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \in V_k(\mathbb{K}^n)$ are given by matrices of the form $\begin{pmatrix} 0 & -T^\dagger \\ T & 0 \end{pmatrix}$, where T is an arbitrary $((n-k) \times n)$ -matrix.

Example 16.3.20 (Canonical connection on the Hopf bundle). Recall that Hopf bundles coincide with the Stiefel bundles $V_1(\mathbb{K}^2) \rightarrow G_1(\mathbb{K}^2)$, where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{H}$. First, consider the complex Hopf bundle. In the notation of the above example, we have

$$u = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{C}^2, \quad |z_1|^2 + |z_2|^2 = 1, \quad (16.3.50)$$

and this the canonical connection is given by

$$\omega^c = \bar{z}_1 dz_1 + \bar{z}_2 dz_2. \quad (16.3.51)$$

It takes values in the Lie algebra $\mathfrak{u}_1 = i\mathbb{R}$ of U_1 . Let us find the local representatives of this connection in terms of the trivializations χ_{\pm} given in Example 7.7.8. The corresponding local sections $s_{\pm} = \chi_{\pm}^{-1}(m, 1) : \mathbb{D}^2 \rightarrow \mathbb{S}^3 \subset \mathbb{C}^2$ read

$$s_+ = \frac{1}{\sqrt{1+|z|^2}}(1, z), \quad s_- = \frac{1}{\sqrt{1+|z|^2}}(z, 1). \quad (16.3.52)$$

By direct substitution we compute

$$\mathcal{A}_+(z) = s_+^* \omega^c = \frac{1}{2} \frac{\bar{z} dz - z d\bar{z}}{1+|z|^2} = -s_-^* \omega^c = -\mathcal{A}_-(z). \quad (16.3.53)$$

In particular, on the equator, $z = e^{i\varphi}$, and the transition function is $\tau_{+-}(z) = z^{-1}$, therefore

$$\mathcal{A}_-(e^{i\varphi}) = -\frac{1}{2} i d\varphi = \mathcal{A}_+(e^{-i\varphi}) - i d\varphi = \mathcal{A}_+(e^{-i\varphi}) + \tau_{+-}^{-1} d\tau_{+-}, \quad (16.3.54)$$

which verifies the transformation law (16.3.34) (recall that $e^{\mp i\varphi}$ represent the same point in charts U_{\pm} , respectively). In spherical coordinates on \mathbb{S}^2 , we have

$$z_+ = \tan \frac{\theta}{2} e^{i\varphi}, \quad z_- = \cot \frac{\theta}{2} e^{-i\varphi}, \quad (16.3.55)$$

and by direct substitution we compute

$$\mathcal{A}_+ = \frac{i}{2} (1 - \cos \theta) d\varphi, \quad \mathcal{A}_- = -\frac{i}{2} (1 + \cos \theta) d\varphi, \quad (16.3.56)$$

and their exterior derivatives coincide (as expected) and define the global representative of the curvature form on \mathbb{S}^2

$$\mathcal{F} = -\frac{dz_{\pm} \wedge d\bar{z}_{\pm}}{1+|z_{\pm}|^2} = \frac{i}{2} \sin \theta d\theta \wedge d\varphi. \quad (16.3.57)$$

Note that this is proportional the standard area form on \mathbb{S}^2 induced from \mathbb{R}^3 , as expected from the SO_3 -invariance of the canonical connection. We will continue exploring connections related to this one in Example 16.4.5.

In contrast with the Hopf bundle, the canonical flat connection on the product bundle $P = \mathbb{S}^2 \times U_1$ in the parametrization $z = e^{i\psi}$ of U_1 reads $\omega = d\psi$.

In complete analogy, for the quaternionic Hopf bundle, we have

$$u = \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix} \in \mathbb{H}^2, \quad |\mathbf{q}_1|^2 + |\mathbf{q}_2|^2 = 1, \quad (16.3.58)$$

and the canonical connection is given by

$$\omega^c = \bar{\mathbf{q}}_1 d\mathbf{q}_1 + \bar{\mathbf{q}}_2 d\mathbf{q}_2. \quad (16.3.59)$$

It takes values in the Lie algebra \mathfrak{sp}_1 of $\text{Sp}_1 = \text{SU}(2)$.

Example 16.3.21 (Canonical connections on a Lie group, part 1). There are two obvious ways of defining a connection on the tangent bundle of a Lie group TG . These connections are induced by the two canonical absolute parallelisms (global trivializations) $TG \rightarrow G \times \mathfrak{g}$ given either by the left or the right Maurer-Cartan forms. The left-invariant (resp. right-invariant) Maurer-Cartan form defines a global frame for TG consisting of left-invariant (resp. right-invariant) vector fields on G . Parallel transport of vectors from g_2 to g_1 can then be defined by left translation $L_{g_1 g_2^{-1}*}$ (resp. $R_{g_1 g_2^{-1}*}$). Clearly both of these connections are flat because the transport is independent of the path.

A natural question is how to realize these connections as principal connections, and whether there are other canonical connections on TG . Recall the construction from Examples 14.2.14 and 14.2.19, where we constructed an isomorphism $\vartheta : TG \rightarrow P \times_G \mathfrak{g}$ between the tangent bundle of a Lie group G and the bundle associated to the trivial principal G -bundle $P = G \times G$. The principal right G -action on P was given by $\Phi_g(g_1, g_2) = (g_1 g, g_2 g)$. The Lie subalgebra corresponding to this action is

$$\mathfrak{h} = \{(A, A) \in \mathfrak{g} \times \mathfrak{g} \mid A \in \mathfrak{g}\}. \quad (16.3.60)$$

There is a continuous family of natural reductive complements parametrized by $\lambda \in \mathbb{R}$:

$$\mathfrak{m}_\lambda = \{((\lambda - 1)A, (\lambda + 1)A) \in \mathfrak{g} \times \mathfrak{g} \mid A \in \mathfrak{g}\}. \quad (16.3.61)$$

It is easy to see that any reductive complement to \mathfrak{h} has to be of this form under the assumption

that $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. The corresponding decomposition of a vector $(X, Y) \in \mathfrak{g} \times \mathfrak{g}$ reads

$$(A, B) = \left(\frac{1+\lambda}{2} A + \frac{1-\lambda}{2} B, \frac{1+\lambda}{2} A + \frac{1-\lambda}{2} B \right) + \left(\frac{\lambda-1}{2} (B-A), \frac{1+\lambda}{2} (B-A) \right). \quad (16.3.62)$$

For each $\lambda \in \mathbb{R}$ we get a principal connection ω^λ on P :

$$\omega^\lambda = \frac{1+\lambda}{2} \text{pr}_1^* \Theta_G + \frac{1-\lambda}{2} \text{pr}_2^* \Theta_G, \quad (16.3.63)$$

where $\text{pr}_1, \text{pr}_2 : G \times G \rightarrow G$ are the two natural projections. Its curvature is computed using the Structure Equation and the Maurer-Cartan equation for Θ_G :

$$\Omega^\lambda = \frac{1-\lambda^2}{4} [\text{pr}_1^* \Theta_G, \text{pr}_2^* \Theta_G]. \quad (16.3.64)$$

The only two flat connections among these are the ones with $\lambda = -1$, and $\lambda = 1$, called ω^- and ω^+ , respectively. We can immediately observe that $\omega^- = \text{pr}_2^* \Theta_G$ and $\omega^+ = \text{pr}_1^* \Theta_G$, so the horizontal distribution for ω^- integrates to the horizontal submanifolds $G \times \{g_2\}$, $g_2 \in G$, and for ω^+ to $\{g_1\} \times G$, $g_1 \in G$. In particular, parallel transports are independent of the path and are given by

$$\text{tra}_{g \sim h}^- : (g_1, g_2) \mapsto (hg^{-1}g_1, g_2), \quad (16.3.65)$$

$$\text{tra}_{g \sim h}^+ : (g_1, g_2) \mapsto (g_1, hg^{-1}g_2), \quad (16.3.66)$$

where $g = g_1 g_2^{-1}$. Note that, projected onto the base G , these parallel transports acts by left translations under ω^- and by right translations under ω^+ . We will continue studying these connections in Example 16.5.18.

16.4 Induced connections

Every principal connection $\mathcal{H}P$ on a principal bundle $P(M, G)$ induces a connection on any associated bundle $E = P^{[\sigma]} = P \times_G F$. For $f \in F$, we define

$$\iota_f : P \rightarrow E, \quad \iota_f(P) = [(p, f)]. \quad (16.4.1)$$

This mapping satisfies

$$\iota_f \circ \Phi_g = \iota_{\sigma_g(f)}, \quad \pi_F \circ \iota_f = \pi. \quad (16.4.2)$$

The horizontal subspace at $e = [(p, f)] \in E$ is defined by

$$\mathcal{H}_e E := (\iota_f)_*(\mathcal{H}_p P). \quad (16.4.3)$$

By (16.4.1), the right hand side of this equation does not depend on the choice of the representative (p, f) . Since $p \mapsto \mathcal{H}_p P$ is a smooth distribution, $e \mapsto \mathcal{H}_e E$ is smooth, too. We show that this distribution is complementary to the canonical vertical distribution $e \mapsto \mathcal{V}_e E$. By the second equation in (16.4.1), $(\iota_f)_*(\mathcal{V}_p P)$ is contained in $\mathcal{V}_e E$ and $\pi_{F*}(\mathcal{H}_e E) = \pi_*(\mathcal{H}_p P) = T_m M$, where $m = \pi_F(e)$. Thus, we have the direct sum decomposition

$$T_e E = \mathcal{V}_e E \oplus \mathcal{H}_e E, \quad (16.4.4)$$

necessary for $\mathcal{H}E$ to be a (general) connection. Note that there is generally no guarantee that this connection is complete. Any vector $X \in T_m M$ has unique horizontal lifts $X_e^h \in T_e E$ and $X_p^h \in T_p P$ related by

$$X_e^h = (\iota_f)_*(X_p^h). \quad (16.4.5)$$

Now consider the special case where $F = V$ is a linear representation of G . Since the map ι_p is a vector space isomorphism between V and $E_{\pi_V(p)}$, its derivative $(\iota_p)_*$ is a linear isomorphism between $T_v V \cong V$ and $\mathcal{V}_e E$, where $e = [(p, v)]$. Thus, for any $Z \in \mathcal{V}_e E$, there exists a vector $v \in V$ such that $Z = (\iota_p)_*(v)$. Via ι_p , the vector v may be identified with the element $[(p, v)] \in E_{\pi_V(p)}$. Thus, the above mentioned identification is given by

$$\mathcal{V}_e E \rightarrow E, \quad Z \mapsto \iota_p \circ (\iota_p)_*^{-1}(Z). \quad (16.4.6)$$

We conclude that in the case of an associated **VB** E , endowed with an induced connection $\mathcal{H}E$, we have the induced connection form $\omega^E \in \Omega^1(E; E)$ given by

$$\omega^E : TE \rightarrow E, \quad \omega^E(X) := \iota_p \circ (\iota_p)_*^{-1}(\mathcal{V}^E(X)), \quad (16.4.7)$$

where $\mathcal{V}^E(X) = X - X^h$ is the vertical projection corresponding to the induced connection $\mathcal{H}E$. Clearly, ω^E is a **VB** morphism. Using $\iota_v \circ \Phi^P = \iota_p \circ \sigma^v$, we find the following relation between ω and ω^E :

$$\omega^E \circ (\iota_v)_*(X) = \iota_p \circ \sigma_*(\omega(X))v, \quad X \in T_p P. \quad (16.4.8)$$

Here, $\sigma_* : \mathfrak{g} \rightarrow \text{End}(V)$ is the representation of the Lie algebra \mathfrak{g} of G induced from σ .

Next, let us discuss the transformation properties of connections under principal bundle morphisms.

Proposition 16.4.1 ([17, Props. 1.3.13, 1.3.15]). *Let (ϑ, λ) be a morphism of the principal bundles $P(M, G)$ and $P'(M', G')$ such that the induced mapping $\vartheta : M \rightarrow M'$ is a diffeomorphism. Let $\mathcal{H}P$ be a principal connection on P and ω its connection form.*

1. *There exists a unique connection $\mathcal{H}P'$ on P' such that ϑ_* maps $\mathcal{H}P$ to $\mathcal{H}P'$.*
2. *The connection form ω' of $\mathcal{H}P'$ fulfills $\vartheta^*\omega' = \lambda_* \circ \omega$, where $\lambda_* : \mathfrak{g} \rightarrow \mathfrak{g}'$ is the induced*

homomorphism of Lie algebras. Moreover, $\vartheta^*\Omega^{\omega'} = \lambda_* \circ \Omega^\omega$.

We call $\mathcal{H}P'$ the image, or pushforward, of $\mathcal{H}P$ under the morphism (ϑ, λ) .

Conversely, if (ϑ, λ) is an arbitrary morphism such that λ is an isomorphism, then a principal connection on P' induces a unique principal connection on P that satisfies the two properties above, called the pullback connection.

Proof. 1. We define the distribution $\mathcal{H}P'$ as follows. Since θ is surjective, for a given $p' \in P'$, we can choose a pair $(p, g) \in P \times G'$ such that $p' = \Phi'_g(\vartheta(p))$ and define

$$\mathcal{H}_{p'}P' := \Phi'_{g*} \circ \vartheta_*(\mathcal{H}_pP). \quad (16.4.9)$$

By equivariance of ϑ , this definition is independent of the choice of (p, g) . We show that $\mathcal{H}P'$ is a connection on P' . First we calculate

$$\Phi'_{h*}(\mathcal{H}_{p'}P') = \Phi'_{h*} \circ \Phi'_{g*} \circ \vartheta_*(\mathcal{H}_pP) = \Phi'_{gh*} \circ \vartheta_*(\mathcal{H}_pP) = \mathcal{H}_{\Phi_h(p')}P', \quad (16.4.10)$$

because $\Phi'_h(p') = \Phi'_{gh}(\vartheta(p))$. Thus, $\mathcal{H}P'$ is G' -equivariant. To prove that it is complementary to VP' , it suffices to show that the restriction of $\pi'_* : T\mathcal{P}' \rightarrow TM'$ to $\mathcal{H}P'$ yields fiberwise isomorphisms. Thus, consider the map $\pi'_* : \mathcal{H}_{p'}P' \rightarrow T_{\pi'(p')}M'$. By G -equivariance of $\mathcal{H}P'$, we may assume $p' = \vartheta(p)$. Then, from $\theta \circ \pi - \pi' \circ \vartheta$, we have

$$\tilde{\vartheta}_{*\pi(p)} \circ \pi_{*p} = \pi_{*p'} \circ \vartheta_{*p}. \quad (16.4.11)$$

Since, by assumption, θ is a diffeomorphism and $\mathcal{H}P$ is a connection, θ_* and π_* are both isomorphisms of vector spaces. Thus, $\pi'_* : \mathcal{H}_{p'}P' \rightarrow T_{\pi'(p')}M'$ is an isomorphism, too. We conclude that $\mathcal{H}P'$ is a connection. By construction, it is unique.

2. The first assertion is equivalent to

$$\omega'_{\vartheta(p)}(\vartheta_*(X)) = \lambda_*(\omega_p(X)), \quad (16.4.12)$$

for any $p \in P$ and $X \in T_pP$. Since ϑ_* maps horizontal vectors to horizontal vectors, it is enough to prove this equality for vertical vectors, that is, for values of Killing vector fields. Thus, let A_* be the Killing vector field generated by $A \in \mathfrak{g}$. Since for any $g \in G$,

$$\omega'_{\vartheta(p)} = \vartheta \circ \Phi_g(p) = \Phi'_{\lambda(g)} \circ \vartheta(p) = \Phi'_{\vartheta(p)} \circ \lambda(g), \quad (16.4.13)$$

we obtain

$$\omega'_{\vartheta(p)}(\vartheta_*(A_*)) = \omega'_{\vartheta(p)}\left(\Phi'_{\vartheta(p)*} \circ \lambda_*(A)\right) = \lambda_*(A). \quad (16.4.14)$$

The assertion now follows from the fact that $A = \omega(A_*)$. The formula for curvature follows immediately via the Structural Equation (16.3.19).

The proof of the statement for pullback connections is very similar. ■

Corollary 16.4.2 ([17, Cor. 1.3.14]). *For a Lie group homomorphism $\lambda : H \rightarrow H$ and a principal H -bundle Q , let $P := Q^{[\lambda]}$. Then any principal connection $\mathcal{H}Q$ induces a unique connection $\mathcal{H}P$ on P . The corresponding connection forms are related via*

$$\vartheta^* \omega^P = \lambda_* \circ \omega^Q, \quad (16.4.15)$$

where ϑ is the corresponding bundle morphism. This induced connection on P is called the λ -extension of ω^Q .

Corollary 16.4.3 ([17, Cor. 1.3.16]). *If $\vartheta : P \rightarrow P'$ is a principal bundle morphism, $G = G'$, and $\lambda = \text{id}_G$, then $\omega = \vartheta^* \omega'$. Thus means, in particular:*

1. *The pullback of a principal connection form under an automorphism of a principal bundle is a principal connection form.*
2. *For a principal bundle $P(M, G)$ and a map $f : N \rightarrow M$, every connection in P induces a connection on the pullback bundle $f^* P$.*

Remark 16.4.4. 1. Proposition 16.4.1 for push-forwards remains true even if θ is any surjective submersion of dimensions of equal dimension. Similarly, its version for pullbacks holds for any λ such that λ_* is an isomorphism of Lie algebras.
 2. From the proof of Proposition 16.4.1, for any morphism $(\vartheta, \lambda) : P_1 \rightarrow P_2$, if $\vartheta^* \omega' = d\lambda \circ \omega$, then $\vartheta^* \Omega^{\omega'} = \lambda_* \circ \Omega^\omega$.
 3. Consider the special case of a fiber product $P_1 \times_M P_2 = \Delta^*(P_1 \times P_2)$, cf. Remark 14.1.10(2). Let $(\pi_i, \lambda_i) : P_1 \times_M P_2 \rightarrow P_i$, $i = 1, 2$, be the natural PFB morphisms defined by restriction of the canonical projections $\text{pr}_i : P_1 \times P_2 \rightarrow P_i$ to $P_1 \times_M P_2$. The corresponding Lie group homomorphisms $\lambda_i : G_1 \times G_2 \rightarrow G_i$ are given by the canonical projections as well. Let ω_i , $i = 1, 2$, be principal connection forms on P_i . Then

$$\omega = \text{pr}_1^* \omega_1 + \text{pr}_2^* \omega_2 \quad (16.4.16)$$

is a connection form on $P_1 \times P_2$. Now, by Corollary 16.4.3, $\vartheta^* \omega$ is the unique connection on $\Delta^*(P_1 \times P_2) = P_1 \times_M P_2$ induced from ω , where $\vartheta : P_1 \times_M P_2 \rightarrow P_1 \times P_2$ is the induced morphism. It is given by

$$\vartheta^* \omega = \pi_1^* \omega_1 + \pi_2^* \omega_2. \quad (16.4.17)$$

Example 16.4.5 (Canonical connection on monopole bundles). Recall the monopole bundles $L(m; 1) \rightarrow \mathbb{S}^2$, $m \in \mathbb{Z}$, from Example 15.2.7. They come with a surjective principal bundle morphism $\vartheta : \mathbb{S}^3 \rightarrow L(m; 1)$ and $\lambda : U_1 \rightarrow U_1$, $z \mapsto z^m$.

Since $\lambda_* : i\mathbb{R} \rightarrow i\mathbb{R}$ is just multiplication by m , the push-forward connection $\omega^{c,m}$ induced on $L(m; 1)$ from the canonical connection ω^c on the Hopf bundle \mathbb{S}^3 (defined in Example 16.3.20)

via Proposition 16.4.1 satisfies

$$\vartheta^* \omega^{c,m} = m\omega^c. \quad (16.4.18)$$

In particular, since ϑ is a vertical morphism, the local representatives \mathcal{A} of this connection on the 2-sphere are just m times those of ω^c :

$$\mathcal{A}_+(z) = \frac{m}{2} \frac{z d\bar{z} - \bar{z} dz}{1 + |z|^2}, \quad \mathcal{A}_-(w) = -\frac{m}{2} \frac{w d\bar{w} - \bar{w} dw}{1 + |w|^2}, \quad (16.4.19)$$

and the (globally defined) curvature is

$$\mathcal{F} = -m \frac{dz \wedge d\bar{z}}{(1 + |z|^2)}. \quad (16.4.20)$$

This is $\frac{im}{2} v_{S^2}$, where $v_{S^2} = \sin \theta d\theta \wedge d\varphi$ is the standard area form on S^2 . Therefore it is easy to compute the total integral of the curvature:

$$\int_{S^2} \mathcal{F} = 2\pi im. \quad (16.4.21)$$

Since the integer m fully classifies all circle bundles over the sphere, this result is pretty surprising in that it relates local connection-dependent information to a global topological invariant of bundles that is independent of a connection. In fact it is even stronger. Let ω' be an arbitrary principal connection on this bundle. Then the difference $\omega' - \omega^{c,m} \in \Omega^1(L(m; 1); g)^{\text{Ad}}$ can be identified with a 1-form $\eta \in \Omega^1(S^2; i\mathbb{R})$. Therefore $\mathcal{F}' - \mathcal{F} = d\eta$ and *the integral of the curvature is independent of the connection*, and is expressed in terms of the only topological invariant that completely classifies U_1 -bundles over S^2 .

In physics, there is an important electromagnetic interpretation of this result. There, $\mathcal{A}_\pm = \frac{ie}{\hbar c} \mathbb{A}_\pm$, where \mathbb{A}_\pm is the magnetic vector potential, and $\mathcal{F} = \frac{ie}{\hbar c} \mathbb{F}$ with \mathbb{F} the magnetic field 2-form. We observe that

$$\mathbb{F} = g \sin \theta d\theta \wedge d\varphi, \quad g = \frac{m\hbar c}{2e}. \quad (16.4.22)$$

Since this is proportional to the area form with coefficient g , this is interpreted as the magnetic field of a point source of strength g , called a *magnetic monopole* of charge g , also called a *Dirac monopole*. As we just showed, *any* U_1 -connection on S^2 must have curvature that integrates to $2\pi im$ for some $m \in \mathbb{Z}$, and thus magnetic charges are *quantized*, namely in units of $\hbar = c = 1$ we must have $2ge \in \mathbb{Z}$. This is our first glimpse of *characteristic classes* (in this case the Chern number, $-m$) whose mathematical role is to classify fiber bundles, playing the role of “quantum numbers” in physics.

Now we consider the very important special case of bundle reductions. Given an injective Lie group homomorphism $\lambda : H \rightarrow G$ and a principal H -bundle Q , we can form the associated principal G -bundle $P = Q^{[\lambda]}$. Then, Q is a λ -reduction of P and P is called a λ -extension of Q . By Corollary 16.4.2, the λ -extension of a connection always exists. The case of a λ -reduction is a bit more involved. Let us assume that (G, H) is a reductive homogeneous space. Then, in general, a principal bundle reduction induces a decomposition of a connection form into a pair of geometrical objects.

Proposition 16.4.6 ([17, Prop. 1.6.8]). *Let $P(M, G)$ be a principal bundle, $H < G$ a closed subgroup and let $Q(M, H)$ be a reduction of P given by the morphism (ϑ, i_H) with $i_H : H \hookrightarrow G$ being the natural inclusion. Assume that the Lie algebra \mathfrak{g} of G admits a reductive decomposition of $\text{Ad}(H)$ -modules $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, with $\mathfrak{h} = \text{Lie } H$. Let ω be a connection form on P and let $\omega_{\mathfrak{h}}, \omega_{\mathfrak{m}}$ be its \mathfrak{h} - and \mathfrak{m} -components, respectively. Then,*

1. $\vartheta^* \omega_{\mathfrak{h}}$ is a connection form on Q ,
2. $\vartheta^* \omega_{\mathfrak{m}} \in \Omega_{\text{hor}}^1(Q; \mathfrak{m})^{\text{Ad}(H)}$.

Proof. Let Φ and Φ' be the G - and H -actions on P and Q , respectively. Then

$$\vartheta \circ \Phi'_h(q) = \Phi_h \circ \vartheta(q), \quad (16.4.23)$$

for any $q \in Q$ and $h \in H$.

1. We must check that $\vartheta^* \omega_{\mathfrak{h}}$ has the properties of a principal connection form. Decomposing $\omega = \omega_{\mathfrak{h}} + \omega_{\mathfrak{m}}$ on P and using that ω is a connection form, for any $A \in \mathfrak{h}$, we have

$$A = \omega(A_*) = \omega_{\mathfrak{h}}(A_*) + \omega_{\mathfrak{m}}(A_*), \quad (16.4.24)$$

where $A_* \in \mathfrak{X}(P)^G$ denotes the Killing vector field generated by A . Thus, $\omega_{\mathfrak{m}}(A_*) = 0$, that is, $\omega_{\mathfrak{h}}(A_*) = A$. Let \tilde{A}_* denote the Killing vector field on Q generated by A . Then, by (16.4.23), $\vartheta_* \circ \tilde{A}_* = A_* \circ \vartheta$ and thus $\vartheta^* \omega_{\mathfrak{h}}(\tilde{A}_*) = A$. It remains to show H -equivariance. For $h \in H$, we have

$$\Phi_h^* \omega = \text{Ad}_h^{-1} \circ \omega = \text{Ad}_h^{-1} \circ \omega_{\mathfrak{h}} + \text{Ad}_h^{-1} \circ \omega_{\mathfrak{m}} \quad (16.4.25)$$

and, on the other hand,

$$\Phi_h^* \omega = \Phi_h^* \omega_{\mathfrak{h}} + \Phi_h^* \omega_{\mathfrak{m}}. \quad (16.4.26)$$

By reductivity, $\text{Ad}_h^{-1} \circ \omega_{\mathfrak{m}}$ takes values in \mathfrak{m} . Hence, $\Phi_h^* \omega_{\mathfrak{h}} = \text{Ad}_h^{-1} \circ \omega_{\mathfrak{h}}$. Then, taking the pullback of this equation under ϑ and using (16.4.23), we obtain the assertion. Moreover, for later use, note that

$$\Phi_h^* \omega_{\mathfrak{m}} = \text{Ad}_h^{-1} \circ \omega_{\mathfrak{m}}. \quad (16.4.27)$$

2. By (16.4.27), $\vartheta^*\omega_m$ is an m -valued 1-form of type $\text{Ad}(H)|_m$. It remains to show that it is horizontal: for any $A \in \mathfrak{h}$, using (16.4.23), we calculate

$$(\vartheta^*\omega_m)_q(\tilde{A}_*) = (\vartheta^*\omega)_q(\tilde{A}_*) - (\vartheta^*\omega_{\mathfrak{h}})(\tilde{A}_*). \quad (16.4.28)$$

The second term yields $-A$. For the first term, we compute

$$(\vartheta^*\omega)_q(\tilde{A}_*) = \omega_{\vartheta(q)}(\vartheta_* \circ \tilde{A}_*(q)) = \omega_{\vartheta(q)}(A_* \circ \vartheta(q)) = A. \quad (16.4.29)$$

■

We conclude that only in the case when $\vartheta^*\omega$ takes values in \mathfrak{h} , its restriction to Q is a connection form on Q . This suggests the following definition.

Definition 16.4.7 (Reducible connection). Let $P(M, G)$ be a principal G -bundle with a principal connection $\mathcal{H}P$ and connection form ω . Let $Q(M, H)$ be a reduction of P given by the morphism (ϑ, i_H) . Then $\mathcal{H}P$ is called reducible to H if $\vartheta^*\omega$ takes values in \mathfrak{h} . $\mathcal{H}P$ is called irreducible if P is not reducible to any proper Lie subgroup of G .

Definition 16.4.8 (Compatible connection). Let $Q(M, H) \subset P(M, G)$ be a principal bundle reduction defined by an element $\varphi \in \text{Hom}_G(P, F)$ taking values in a single orbit O and by a point $f \in O$. A principal connection $\mathcal{H}P$ on P is called compatible with φ if it is reducible to Q .

16.5 Holonomy and reduction

In this section we explore the fundamental relationship between bundle reductions and holonomy. The basic question that we would like to answer is: what is the “smallest possible” reduction of a given principal bundle with a principal connection? Recall from Section 16.3 a principal connection ω defines a global parallel transport functor on the principal bundle P which we denote by tra^ω . In particular, if $\gamma \in C(m)$ is a loop based at $m \in M$, then $\text{tra}_\gamma^\omega : P_m \rightarrow P_m$ is a diffeomorphism. It is equivariant under the principal right G -action since if γ^h is the horizontal lift of γ with starting point $p \in P_m$, then obviously $\gamma^h \cdot g = \Phi_g \circ \gamma^h$ is the horizontal lift starting at $p \cdot g = \Phi_g(p)$. Thus

$$\text{tra}_\gamma^\omega(p \cdot g) = \text{tra}_\gamma^\omega(p) \cdot g. \quad (16.5.1)$$

In addition, since the G -action is free and transitive on P_m , there exists a unique $g_\gamma \in G$ such that

$$\text{tra}_\gamma^\omega(p) = \gamma^h(1) = p \cdot g_\gamma. \quad (16.5.2)$$

It is easy to see that

$$\text{tra}_\gamma^\omega \circ \text{tra}_{\gamma'}^\omega = \text{tra}_{\gamma \bullet \gamma'}^\omega, \quad (16.5.3)$$

so that tra^ω , given a basepoint $p \in P$, defines an anti-homomorphism $C(m) \rightarrow G$ given by $\gamma \mapsto g_\gamma$ (it would have been a homomorphism if the principal action Φ were a left action). We can therefore also define the holonomy groups based at $p \in P$ as subgroups of G :

$$\text{Hol}_p(\omega) := \{g_\gamma \mid \gamma \in C(m)\}, \quad \text{Hol}_p^0(\omega) := \{g_\gamma \mid \gamma \in C^0(m)\}. \quad (16.5.4)$$

Obviously $\text{Hol}_p(\omega)$ and $\text{Hol}_m(\omega)$ are isomorphic as abstract groups, although there is no natural isomorphism since we have to pick a basepoint $p \in P_m$.

Remark 16.5.1. Consider the following equivalence relation on P : $p_1 \sim p_2$ iff they can be joined by a ω -horizontal curve. Then $\text{Hol}_p(\omega)$ coincides with the subset of elements $g \in G$ such that $p \sim p \cdot g$.

The following proposition is trivial.

Proposition 16.5.2 ([17, Prop. 1.7.8]). *Let ω be a principal connection form on a (right) principal bundle $P(M, G)$.*

1. *The holonomy groups of ω with base points p and $p \cdot g$, $g \in G$, are conjugate:*

$$\text{Hol}_{p \cdot g}(\omega) = g^{-1} \text{Hol}_p(\omega) g. \quad (16.5.5)$$

(replace $g \rightarrow g^{-1}$ on the right in case of a left principal bundle).

2. *If two points in P can be joined by a horizontal curve, then their holonomy groups coincide.*

If M is connected, then for each pair $p_1, p_2 \in P$ there exists a group element $g \in G$ such that p_1 and $p_2 \cdot g$ can be joined by a horizontal curve. In this case all holonomy groups $\text{Hol}_p(\omega)$, $p \in P$, are conjugate in G .

Theorem 16.5.3 ([17, Thm. 1.7.9]). *Let $P(M, G)$ be a principal bundle over a connected manifold M and let ω be a principal connection form on it. Then, for every $p \in P$,*

1. $\text{Hol}_p^0(\omega)$ is a connected Lie subgroup of G ,
2. $\text{Hol}_p^0(\omega)$ is a normal subgroup of $\text{Hol}_p(\omega)$ and $\text{Hol}_p(\omega)/\text{Hol}_p^0(\omega)$ is countable.

Proof. Recall that we proved Yamabe's Theorem 12.7.1 with the additional assumption that the subgroup is C^∞ -path-connected. We now prove that the restricted holonomy group is C^∞ -path-connected and apply Yamabe's Theorem.

Let $m = \pi(p)$, let $[0, 1] \ni s \mapsto \gamma(s) \in M$ be an element of $C^0(m)$ and let γ^h be its horizontal lift starting at p . Choose a smooth homotopy $\tau : [0, 1]^2 \rightarrow M$ written as γ_t , $t \in [0, 1]$, such that

$$\tau(1, s) = \gamma, \quad \tau(0, s) = \tau(t, 0) = \tau(t, 1) = m, \quad (16.5.6)$$

for all $(t, s) \in [0, 1]^2$.

By Corollary 16.4.3, the connection ω induces a connection ω^τ on the pullback bundle τ^*P . Let $\vartheta : \tau^*P \rightarrow P$ be the induced morphism projection to τ . Clearly, for every $t \in [0, 1]$, the preimage under τ of the closed curve $s \mapsto \tau(t, s)$ is the line segment $s \mapsto (t, s)$ in \mathbb{R}^2 . Let F_t be the flow of the ω^τ -horizontal lift of ∂_s . Then

$$t \mapsto \vartheta \circ F_t(1) \quad (16.5.7)$$

is a smooth curve in P starting at p_0 and ending at $\gamma^h(1)$. Via (16.5.2), it defines a smooth curve in G starting at the identity and ending at $g_\gamma \in \text{Hol}_{p_0}^0(\omega)$, that is, $\text{Hol}_{p_0}^0(\omega)$ is a C^∞ -path-connected subgroup of G , and hence a Lie subgroup by Yamabe's Theorem.

Now let us prove the second assertion. Clearly, for smooth closed curves γ, γ' starting at $m_0 \in M$, with γ nullhomotopic, the curve $\gamma' \bullet \gamma \bullet \gamma'^{-1}$ is nullhomotopic, too. Thus $\text{Hol}_p^0(\omega)$ is a normal subgroup of $\text{Hol}_p(\omega)$. To prove that the quotient group is countable, we define a homomorphism $F : \pi_1(M, m) \rightarrow \text{Hol}_p(\omega)/\text{Hol}_p^0(\omega)$ as follows. For a given $\alpha \in \pi_1(M, m)$, let $t \mapsto \gamma(t)$ be a smooth curve representing α (it exists e.g. by Whitney approximation). We put $F[\alpha] := [g_\gamma]$. This map is well-defined because a different choice of $\gamma \in [\alpha]$ differs by a nullhomotopic path, this the corresponding g_γ will change by a conjugation by an element of $\text{Hol}_p^0(\omega)$. Clearly F is surjective. Thus, F is a (anti)-homomorphism onto $\mathcal{H}_p(\omega)/\text{Hol}_p^0(\omega)$. Now, countability of this quotient follows from the countability of $\pi_1(M, m)$ (because there is an everywhere dense countable subset in M , and each loop can be made homotopic to a product of pre-chosen paths connecting pairs of points of that set; there are countably many such paths). ■

Remark 16.5.4. Thus $\text{Hol}_p(\omega)$ is a Lie subgroup of G whose identity component is $\text{Hol}_p^0(\omega)$. In particular, if M is simply connected, then $\text{Hol}_p(\omega)$ is connected.

Proposition 16.5.5 ([17, Prop. 1.7.12]). *Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a principal bundle with a connected base M and let ω be a principal connection form on P . Let $p_0 \in P$ and let $P_{p_0}(\omega)$ be the subset of points in P which can be joined to p_0 by a ω -horizontal path. Then*

1. *$P_{p_0}(\omega)$ is a reduction of P with structure group $\text{Hol}_{p_0}(\omega)$.*
2. *The connection ω is reducible to a principal connection on $P_{p_0}(\omega)$.*

Proof. 1. Since M is connected, the restriction of π to $P_{p_0}(\omega)$ is surjective. By Proposition 16.5.2 and Remark 16.5.1, $P_{p_0}(\omega)$ is invariant under the right action of $\text{Hol}_{p_0}(\omega) \subset G$ and $P_{p_0}(\omega)$ intersects the fibers of P in $\text{Hol}_{p_0}(\omega)$ -orbits.

Next we show that $P_{p_0}(\omega)$ is a subbundle of P . For that purpose, let $p \in P_{p_0}(\omega)$ and let (U, κ) be a local chart for M at $m = \pi(p)$ such that $\kappa(U) \subset \mathbb{R}^{\dim M}$ is an open ball and

$\kappa(m) = 0$. For any $\tilde{m} \in U$, let $t \mapsto \gamma(t)$ be the unique curve $m \rightsquigarrow \tilde{m}$ such that $t \mapsto \kappa \circ \gamma(t)$ is the line segment $0 \rightsquigarrow \kappa(\tilde{m})$. Define

$$s : U \rightarrow P, \quad s(\tilde{m}) := g_\gamma(p). \quad (16.5.8)$$

Clearly, s is a smooth local section fulfilling $s(U) \subset P_{p_0}(\omega)$. Now, for every $p \in \pi^{-1}(U)$, there exists a unique $g \in G$ such that $p = s(\pi(p)) \cdot g$. Then

$$\tilde{\chi} : \pi^{-1}(U) \rightarrow U \times G, \quad \tilde{\chi}(p) := (\pi(p), g) \quad (16.5.9)$$

is a bijective map which induces a bijective map

$$\chi : P_{p_0}(\omega) \cap \pi^{-1}(U) \rightarrow U \times \text{Hol}_{p_0}(\omega). \quad (16.5.10)$$

Constructing in this way a system of bijective mappings $\{(U_\alpha, \kappa_\alpha)\}$ such that $\{U_\alpha\}$ is a covering of M and requiring that the maps χ_α are diffeomorphisms, we endow $P_{p_0}(\omega)$ with a manifold structure and a system of local trivializations. To see that it is a submanifold of P , note that $\{\pi^{-1}(U_\alpha)\}$ is a covering of $P_{p_0}(\omega)$ with open subsets of P such that every subset $P_{p_0}(\omega) \cap \pi^{-1}(U)$ is a submanifold of P . This follows from the fact that $\text{Hol}_{p_0}(\omega)$ is a submanifold of G and that the χ_α are diffeomorphisms.

We conclude that $P_{p_0}(\omega)$ is a reduction of P with structure group $\text{Hol}_{p_0}(\omega)$ with the corresponding morphism (ϑ, λ) given by the natural inclusion maps $\vartheta : P_{p_0}(\omega) \rightarrow P$ and $\lambda : \text{Hol}_{p_0}(\omega) \rightarrow G$.

2. Let $p \in P_{p_0}(\omega)$ and let $X \in \mathcal{H}_p P$. Then there exists a curve γ in M starting at $\pi(p)$ such that the horizontal lift γ^h starting at p fulfills $X = \dot{\gamma}^h(0)$. Since p can be joined to p_0 by a horizontal path, the image of γ^h is contained in $P_{p_0}(\omega)$ and $X \in T_p(P_{p_0}(\omega))$. Thus, for every $p \in P_{p_0}(\omega)$, the horizontal subspace $\mathcal{H}_p P$ is tangent to $P_{p_0}(\omega)$. This means that the connection ω (or $\mathcal{H}P$) is reducible to $P_{p_0}(\omega)$ the horizontal subspace at $p \in P_{p_0}(\omega)$ of the reduced connection is given by $\mathcal{H}_p P$. ■

Definition 16.5.6 (Holonomy bundle). The subbundle $P_{p_0}(\omega)$ constructed above is called the holonomy bundle of ω with base point p_0 .

Note that $P_{p_0}(\omega) = P_{p_1}(\omega)$ iff p_0, p_1 can be joined by a horizontal curve. Thus, for any pair $p_0, p_1 \in P$, we have either $P_{p_0}(\omega) = P_{p_1}(\omega)$ or $P_{p_0}(\omega) \cap P_{p_1}(\omega) = \emptyset$, that is, P is foliated by disjoint holonomy bundles. One can check that all holonomy bundles of a given connection are isomorphic.

Remark 16.5.7. Let $P(M, G)$ be a principal bundle, let $H < G$ be a Lie subgroup and let $Q(M, H)$ be a reduction of P given by a morphism (ϑ, i_H) with $i_H : H \hookrightarrow G$ the natural inclusion. Let $\mathcal{H}P$ be a principal connection on P that is reducible to a principal connection

$\mathcal{H}Q$. Then, by Proposition 16.3.28, $\mathcal{H}Q$ defines a connection $\hat{\mathcal{H}}P$ on P (the image of $\mathcal{H}Q$ under ϑ). Now, $\hat{\mathcal{H}}P$ is either irreducible or not. In the first case, $\vartheta(Q)$ coincides with the holonomy bundle $P_p(\mathcal{H}P)$, $p \in \vartheta(Q)$, and $\hat{\mathcal{H}}P$ coincides with the reduction of $\mathcal{H}P$ to $P_p(\mathcal{H}P)$. In the second case, by Proposition 16.5.5, $\hat{\mathcal{H}}P$ is reducible to the holonomy bundle. Thus, in this case, for all $p \in \vartheta(Q)$, we have

$$P_p(\mathcal{H}P) \subset \vartheta(Q), \quad \hat{\mathcal{H}}P|_{P_p(\mathcal{H}P)} = \mathcal{H}P|_{P_p(\mathcal{H}P)}. \quad (16.5.11)$$

Thus the holonomy bundle is the smallest possible reduction of a principal bundle with connection. In particular, a principal connection $\mathcal{H}P$ is irreducible iff $P = P_p(\mathcal{H}P)$ and $G = \text{Hol}_p(\mathcal{H}P)$ for all $p \in P$.

The key consequence of this result is that we have reduced the study of bundle reductions to the study of the affine space of connections $\mathcal{A}(P)$. In this way the classification of bundles becomes a problem of linear (in fact homological) algebra.

The following classical theorem characterizes the Lie algebra of the holonomy group of a connection in terms of its curvature. The Lie algebra turns out to consist precisely of the values of the curvature form on horizontal vectors along the holonomy bundle.

Theorem 16.5.8 (Ambrose-Singer [17, Thm. 1.7.15]). *Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a principal bundle with connected base M and let $\mathcal{H}P$ be a connection on P with connection form ω and curvature form Ω^ω . Let $\mathfrak{g} = \text{Lie } G$ be the Lie algebra of G . Then, for any $p_0 \in P$, the Lie algebra \mathfrak{hol}_{p_0} of the holonomy group $\text{Hol}_{p_0}(\omega) = \text{Hol}_{p_0}(\mathcal{H}P)$ coincides with the subspace of \mathfrak{g} generated by elements of the form $\Omega_p(X, Y)$, where $p \in P_{p_0}(\omega)$ and $X, Y \in \mathcal{H}_p P$.*

Proof. By Proposition 16.5.5, WLOG we may assume $\text{Hol}_{p_0}(\omega) = G$ and $P_{p_0}(\omega) = P$. Let

$$\mathfrak{hol} := \text{span} \{ \Omega_p^\omega(X, Y) \in \mathfrak{g} : p \in P_{p_0}(\omega), X, Y \in \mathcal{H}_p P \}. \quad (16.5.12)$$

We must show that $\mathfrak{hol} = \mathfrak{g}$. First, since Ω^ω is a horizontal form of type Ad, the subspace \mathfrak{hol} is Ad(G)-invariant. Thus \mathfrak{hol} is an ideal in \mathfrak{g} . Next, consider the distribution

$$p \mapsto D_p := \mathcal{H}_p P \oplus \Phi_*^P(\mathfrak{hol}). \quad (16.5.13)$$

We show that D is involutive. Since $\mathcal{H}_p P$ is spanned by horizontal vector fields and $\Phi_*^P(\mathfrak{hol})$ is spanned by Killing vector fields generated by elements of \mathfrak{hol} , we must consider the following three cases:

- (a) Let $A, B \in \mathfrak{hol}$ and let A_*, B_* be the corresponding Killing vector fields. Then $[A, B] \in \mathfrak{hol}$ and since $[A_*, B_*] = [A, B]_*$, we have $[A_*, B_*]_p \in \Phi_*^P(\mathfrak{hol})$.

- (b) Let X be a horizontal vector field and let $A \in \mathfrak{hol}$. Then, by Lemma 16.3.10, $[X, A_*]$ is a horizontal vector field.
- (c) Let X and Y be horizontal vector fields. Then, by definition of the curvature form,

$$\mathcal{V}([X, Y]_P) = -\Phi_*^P(\Omega^\omega(X, Y)). \quad (16.5.14)$$

Now, the Frobenius Theorem 9.5.4 and Theorem 9.5.11 yield the existence of a maximal connected integral manifold N through $p_0 \in P$. A point $p \in P$ belongs to N iff there exists a curve $\gamma : p \rightsquigarrow p_0$ such that $\dot{\gamma}(t) \in D_{\gamma(t)}$ for all t . Since $\mathcal{H}P \subset TN$, we conclude $P_{p_0}(\omega) = P \subset N$. Thus $N = P$ and $D = TP$ amnd, consequently,

$$\dim \mathfrak{g} = \dim P - \dim M = \dim N - \dim M = \dim \mathfrak{hol}, \quad (16.5.15)$$

that is, $\mathfrak{g} = \mathfrak{hol}$. ■

Remark 16.5.9. 1. As we saw, P is foliated by holonomy bundles. By the proof of the Ambrose-Singer Theorem, this is exactly the foliation of the distribution D .

- 2. If the curvature Ω^ω vanishes, then $\text{Hol}_p^0(\omega) = \{e\}$ and each holonomy bundle $P_p(\omega)$ is isomorphic to the trivial bundle $M \times G$ and the isomorphism maps $\mathcal{H}P$ to the canonical connection on $M \times G$, cf. Example 16.3.16.
- 3. If Ω^ω vanishes and if, additionally, M is simply connected, then P is isomorphic to the trivial bundle $M \times G$ and the isomorphism again maps $\mathcal{H}P$ to the canonical connection on $M \times G$.
- 4. If G is connected, then $\mathcal{H}P$ is irreducible iff $\mathfrak{hol}_{p_0}(\omega) = \mathfrak{g}$.
- 5. Using the Ambrose-Singer Theorem, one can show that every PFB over a base M with $\dim M \geq 2$ admits an irreducible principal connection.

Let us now examine parallel transport on associated bundles. First we note that parallel transport on principal bundles $P(M, G)$ translates over to associated bundles $E = P \times_G F$. The unique lift of a path γ in M to the horizontal path γ_E^h in E starting at the point $[(p, f)] \in E_{\gamma(0)}$ is given by

$$\gamma_E^h(t) = \iota_f(\gamma^h(t)), \quad (16.5.16)$$

where γ^h is the horizontal lift to P starting at p and $\iota_f : P \rightarrow E$ was defined in (16.4.1). The corresponding parallel transport operators along γ will be denoted by

$$\text{tra}_\gamma^{\mathcal{H}E}(t) : E_{\gamma(0)} \rightarrow E_{\gamma(t)}, \quad (16.5.17)$$

and in the case $t = 1$ we simply write $\text{tra}_\gamma^{\omega^E}$ as usual. For a given lift γ^h to P , formula (16.5.1) implies

$$\text{tra}_\gamma^{\mathcal{H}E}(t) = \iota_{\gamma^h(t)} \circ (\iota_{\gamma^h(0)})^{-1}. \quad (16.5.18)$$

It is easy to check that this is indeed the parallel transport functor corresponding to the induced connection $\mathcal{H}E$. Therefore we have shown

Corollary 16.5.10. *If a general connection $\mathcal{H}E$ on a bundle E associated to a principal bundle P is induced from a principal connection on P , then it is complete. In addition, parallel transport acts by isomorphisms of fiber G -structures.*

Remark 16.5.11. Recall that it was originally unclear how to define connections on arbitrary bundles in a way that respects any present G -structure. The only meaningful definition would have been to say that the local Christoffel forms Γ^α take values in the set of Killing vector fields of the G -action σ on the typical fiber F . We have worked around this issue by inducing such connections from principal connections. Such connections automatically turn out to be complete.

It is easy to see that the converse holds as well. In fact, provided the G -action on the typical fiber F is infinitesimally faithful (i.e. $\sigma_* : \mathfrak{g} \rightarrow \mathfrak{X}(F)$ is injective), every general connection whose Christoffel symbols are $\sigma_*(G)$ -valued is induced from a principal connection [43]. Indeed, such Christoffel forms uniquely define \mathfrak{g} -valued local forms $\mathcal{A}_\alpha \in \Omega^1(U_\alpha; \mathfrak{g})$ via $\Gamma^\alpha(X) = \sigma_*(\mathcal{A}_\alpha(X))$. The transformation law for \mathcal{A}_α also follows from that for Γ^α and hence we get a uniquely determined principal connection form.

We have therefore completed the description of the category of connections on bundles with G -structures.

In particular, for loops in M , we can define the holonomy group

$$\text{Hol}_m(\mathcal{H}E) := \{\text{tra}_\gamma^{\mathcal{H}E} : \gamma \in C(m)\} \subset \text{Diff}(E_{\gamma(0)}), \quad (16.5.19)$$

and, correspondingly the restricted holonomy group $\text{Hol}_m^0(\mathcal{H}E)$. Using (16.5.16) it is easy to show that for any $p \in P_m$,

$$\text{Hol}_m(\mathcal{H}E) = \iota_p \circ \sigma(\text{Hol}_p(\omega)) \circ \iota_p^{-1}. \quad (16.5.20)$$

In particular, if σ is injective (faithful), then $\text{Hol}_m(\mathcal{H}E)$ and $\text{Hol}_p(\omega)$ are isomorphic.

Definition 16.5.12 (Parallel section). A section $s \in \Gamma^\infty(E)$ is called parallel with respect to a given connection $\mathcal{H}E$ if $\text{im}(s_{*m}) \subset \mathcal{H}_{s(m)} E$ for all $m \in M$.

Proposition 16.5.13 ([17, Prop. 1.5.6]). *Let $s \in \Gamma^\infty(E)$ be a section and $\tilde{s} \in \text{Hom}_G(P, F)$ be the associated equivariant function. Then for any $X \in \mathfrak{X}(M)$,*

$$\mathcal{V}^E(s_*(X)) = (\iota_p)_* \circ X^h(\tilde{s}). \quad (16.5.21)$$

Proof. By the definition of \mathcal{V}^E , we must decompose $s_*(X)$ into its vertical and horizontal parts. For that purpose, let $t \mapsto \gamma(t)$ be an integral curve of X through $m \in M$. Then

$$s_*(X_m) = \frac{d}{dt} \Big|_0 s \circ \gamma(t). \quad (16.5.22)$$

Choose a point $p \in P_m$ and take the integral curve $t \mapsto \gamma^h(t)$ through p of the horizontal lift X^h of X to P . Then (14.2.36) implies

$$s \circ \gamma(t) = \iota \left(\gamma^h(t), \tilde{s}(\gamma^h(t)) \right) \quad (16.5.23)$$

and thus

$$\begin{aligned} \frac{d}{dt} \Big|_0 s \circ \gamma(t) &= \frac{d}{dt} \Big|_0 \iota \left(\gamma^h(t), \tilde{s}(\gamma^h(t)) \right) = \\ &= \frac{d}{dt} \Big|_0 \iota_{\tilde{s}(p)}(\gamma^h(t)) + \frac{d}{dt} \Big|_0 \iota_p \left(\tilde{s}(\gamma^h(t)) \right). \end{aligned} \quad (16.5.24)$$

For the first term, using (16.4.5), we have

$$\frac{d}{dt} \Big|_0 \iota_{\tilde{s}(p)}(\gamma^h(t)) = (\iota_{\tilde{s}(p)})_* X_p^h = X_{s(m)}^h. \quad (16.5.25)$$

This is the horizontal part of $s_*(m)$ at $s(m)$. The second term reads

$$\frac{d}{dt} \Big|_0 \iota_p \left(\tilde{s}(\gamma^h(t)) \right) = (\iota_p)_p \left(\frac{d}{dt} \Big|_0 \tilde{s}(\gamma^h(t)) \right) = (\iota_p)_*(X^h(\tilde{s})). \quad (16.5.26)$$

This is the vertical component of $s_*(X) \in T_{s(m)} E$. Thus, by defintion of \mathcal{V}^E , the assertion follows. ■

Remark 16.5.14. From the above proof, $X^h(s(m))$ is the horizontal part of $s_*(X)$. Therefore the vertical part can also be written as $\mathcal{V}^E(s_*(X)) = s_* \circ X - X^h \circ s$. This object quantifies the failure of s to be parallel. We can denote it by a separate symbol:

$$\nabla_X^{\mathcal{V}} s := s_* \circ X - X^h \circ s = \mathcal{V}^E(s_*(X)) = (\iota_p)_* \circ X^h(\tilde{s}) \in \Gamma^\infty(\mathcal{V}E). \quad (16.5.27)$$

This is called the *absolute derivative* of s along X . It takes $s \in \Gamma^\infty(E)$ and $X \in \mathfrak{X}(M)$ and produces $\nabla_X s \in \Gamma^\infty(\mathcal{V}E)$. The general formula relating Lie derivatives to flows also implies

$$\nabla_X^{\mathcal{V}} s = \frac{d}{dt} \Big|_0 \left(F_{-t}^{X^h} \circ s \circ F_t^X \right). \quad (16.5.28)$$

At this point this object isn't very useful because it takes values in a bundle that is very different from the original one. The situation is better on principal and vector bundles, where we can replace \mathcal{V}^E by either ω or ω^E . We will treat the case of **VBs** later, but on principal bundles we can redefine

$$\nabla_X^\omega s := \omega(s_*(X)), \quad s \in \Gamma^\infty(P), X \in \mathfrak{X}(M), \quad (16.5.29)$$

so that $\nabla s \in \Omega^1(M; \mathfrak{g})$.

Example 16.5.15. We continue Example 16.3.16 of the canonical connection $\omega = \text{pr}_G^* \Theta_G$ on the product bundle $P = M \times G \rightarrow M$. In this bundle, sections $s : M \rightarrow P$ can be viewed as functions $s : M \rightarrow G$, and the absolute derivative $\nabla^\omega s$ coincides with the left logarithmic derivative $\delta s \in \Omega^1(M; \mathfrak{g})$.

Corollary 16.5.16. *In particular, s is parallel with respect to $\mathcal{H}E$ iff*

$$X^h(\tilde{s}) = 0 \text{ for all } X^h \in \mathcal{H}P. \quad (16.5.30)$$

Denote the subspace of sections of E that are parallel with respect to $\mathcal{H}E$ by $\mathcal{P}(E, \mathcal{H}E)$. We now show that the holonomy group of ω^E can be used to characterize the set $\mathcal{P}(E, \mathcal{H}E)$ of parallel sections. Given $p \in P$, and element of F is called holonomy-invariant if it is invariant under the restriction of the action σ to the holonomy group $\text{Hol}_p(\omega)$.

Proposition 16.5.17 (Holonomy principle [17, Prop. 1.7.20]). *If M is connected, then there is a bijection between $\mathcal{P}(E, \mathcal{H}E)$ and the space of holonomy-invariant points in F .*

Proof. Let $m_0 \in M$ and $p_0 \in P_{m_0}$. By Proposition 16.5.5, P reduces together with ω to the holonomy bundle $P_{p_0}(\omega)$ and, by Proposition 14.4.8, we have the following isomorphism of associated **VBs**:

$$E = P \times_G F \cong P_{p_0}(\omega) \times_{\text{Hol}_{p_0}(\omega)} F. \quad (16.5.31)$$

Thus, it is enough to consider sections of the associated bundle on the right hand side which are parallel in the sense of the reduced connection.

1. Let $f \in F$ be holonomy-invariant, that is, $\sigma_h(f) = f$ for all $h \in \text{Hol}_{p_0}(\omega)$. Define

$$\tilde{s} : P_{p_0}(\omega) \rightarrow F, \quad \tilde{s}(p) := f. \quad (16.5.32)$$

Since, for all $h \in \text{Hol}_{p_0}(\omega)$, we have

$$\tilde{s}(p \cdot h) = f = \sigma_{h^{-1}}f = \sigma_{h^{-1}}\tilde{s}(p), \quad (16.5.33)$$

\tilde{s} is $\text{Hol}_{p_0}(\omega)$ -equivariant. Thus, by Proposition 14.2.17, it induces a smooth section s of E . Since \tilde{s} is constant on $P_{p_0}(\omega)$, we have $X^h(\tilde{s}) = X^h(f) = 0$ for any $X \in \mathfrak{X}(M)$. Then Proposition 16.5.13 implies that s is parallel.

2. Conversely, let s be a parallel section and let \tilde{s} be the corresponding equivariant function. By Proposition 16.5.13, we have $X^h(\tilde{s}) = 0$ for every horizontal vector field on $P_{p_0}(\omega)$. Thus \tilde{s} is constant along any horizontal curve in $P_{p_0}(\omega)$, that is, \tilde{s} is constant on $P_{p_0}(\omega)$. Let $f := \tilde{s}(p) \in F$ be this constant value. The equivariance of \tilde{s} implies the holonomy-invariance of f .

■

Example 16.5.18 (Canonical connections on a Lie group, part 2). We continue Example 16.3.21. Clearly $\omega^{\pm 1}$ are the only two flat canonical connections, so they have trivial holonomy. Conversely, for $\lambda \neq \pm 1$, and assuming G is simple (in particular, $[G, G] = G$), the values of Ω^λ span all of \mathfrak{g} , and therefore, by the Ambrose-Singer Theorem, all holonomy groups are all of G , so the connections are irreducible.

For each λ we get an induced linear connection ∇^λ on the associated bundle $P \times_G \mathfrak{g}$, where G acts on \mathfrak{g} via the adjoint representation. Via the isomorphism $\vartheta : TG \rightarrow P \times_G \mathfrak{g}$ defined in Example 14.2.14, we identify ∇^λ with a connection on TG . We would like to characterize parallel vector fields under this connection. Note that, by the Holonomy Principle, such vector fields correspond to holonomy-invariant elements of \mathfrak{g} , and under the assumption that $[G, G] = G$, nonzero holonomy-invariant elements exist only if $\lambda = \pm 1$, and the connection is irreducible. Therefore we focus on these two connections, called ∇^+ and ∇^- .

If $\lambda = \pm 1$, then the holonomy is trivial, and the holonomy bundle of any point $p_0 = (g_1, g_2) \in P$ is either $P_{p_0}^- = \{(gg_1, g_2) \mid g \in G\}$ or $P_{p_0}^+ = \{(g_1, gg_2) \mid g \in G\}$. The reduced connection form on $P_{p_0}^\pm$ is the restriction of ω^\pm , which is readily confirmed to be 0, as expected from the fact that the vertical bundle of $P_{p_0}^\pm$ has rank 0. By Proposition 14.4.8, $P_{p_0}^\pm$ induces an isomorphism between $P \times_G \mathfrak{g}$ and $P_{p_0}^\pm \times \mathfrak{g}$, with the latter being the product bundle (since the structure group of $P_{p_0}^\pm$ is $\{e\}$). Since $P_{p_0}^\pm$ is diffeomorphic to G itself, it is easy to see that these two isomorphisms correspond exactly to the two canonical global trivializations of TG : for $\lambda = -1$, it is the left-invariant trivialization $(g, X) \mapsto (g, L_g^{-1}X)$, and for $\lambda = 1$ – the right-invariant one, $(g, X) \mapsto (g, R_{g_*}^{-1}X)$.

By the Holonomy Principle, parallel vector fields are in one-to-one correspondence with elements of \mathfrak{g} , and in particular $A \in \mathfrak{g}$ corresponds to the constant function on $P_{p_0}^\pm$ with value A . Under the isomorphisms described above, combined with ϑ , these are exactly the left-invariant vector fields for ∇^- , and the right-invariant ones for ∇^+ . This agrees with our finding in Example 16.3.21 that parallel transports under ω^- are left translations, and under ω^+ right translations. Therefore these two connections on TG coincide with those we considered in Example 16.1.8. We will continue studying these connections in Example 16.8.15.

16.6 Gauge transformations

Let us discuss the connection to the gauge group. The Lie algebra of the group of automorphisms (not only vertical ones) $\text{Aut}(P)$ is the space $\text{aut}(P) = \mathfrak{X}(P)^G$ of G -invariant vector fields on P , and the Lie algebra of the gauge group (of vertical automorphisms) is the subspace $\text{gau}(P)$ of vertical G -invariant vector fields. One has the exact sequence of Lie algebras

$$0 \rightarrow \text{gau}(P) \rightarrow \text{aut}(P) \rightarrow \mathfrak{X}(M) \rightarrow 0. \quad (16.6.1)$$

The surjectivity of the map $\text{aut}(P) \rightarrow \mathfrak{X}(M)$ can be proved using a trivializing open covering of M and a **POU** subordinate to that cover.

For $\varphi \in \text{Hom}_G(P, G)$, the corresponding vertical automorphism is given by

$$\vartheta_\varphi : P \rightarrow P, \quad p \mapsto p \cdot \varphi(g). \quad (16.6.2)$$

The equivariance property $\vartheta_\varphi(p \cdot g) = \vartheta_\varphi(p) \cdot g$ is clearly equivalent to $\varphi(p \cdot g) = g^{-1}\varphi(p)g$, which is exactly the defining property of $\text{Hom}_G(P, G)$. It is also easy to check that this correspondence is a group isomorphism (where elements of $\text{Hom}_G(P, G)$ are multiplied pointwise, so $\vartheta_\varphi \circ \vartheta_\psi = \vartheta_{\varphi \cdot \psi}$). Finally, recall from Proposition 14.2.17 that the space $\text{Hom}_G(P, G)$ can be identified with the space of sections of the associated group bundle $\mathbf{Ad}(P) = P \times_G G$, where G acts on itself by conjugations. We thus have the isomorphisms

$$\text{Aut}(P) \cong \text{Hom}_G(P, G) \cong \Gamma^\infty(\mathbf{Ad}(P)). \quad (16.6.3)$$

It is also obvious that every vertical automorphism $\vartheta_\varphi \in \text{Gau}(P)$ induces a vertical automorphism $\hat{\vartheta}_\varphi$ of any associated bundle $P^{[\sigma]} = P \times_G F$ given by

$$\hat{\vartheta}_\varphi([(p, f)]) = [(p, \sigma_{\varphi(p)}(f))]. \quad (16.6.4)$$

In particular, if $G \xrightarrow{\sigma} F$ is a linear representation, then $\hat{\vartheta}_\varphi$ is a vertical automorphism of vector bundles.

Proposition 16.6.1 ([17, Prop. 1.8.7]). *Let $(P \xrightarrow{\pi} M, G, \Phi)$ be a principal bundle and let $\mathcal{H}P$ be a principal connection. Then, for $\vartheta \in \text{Gau}(P)$ corresponding to $\varphi \in \text{Hom}_G(P, G)$, one has the following transformation laws:*

1. *If ω is the connection form of $\mathcal{H}P$, then $\vartheta^*\omega$ is the connection form of $\vartheta_*^{-1}(\mathcal{H}P)$ and*

$$(\vartheta^*\omega)_p = \text{Ad}_{\varphi(p)}^{-1} \circ \omega_p + (\varphi^*\Theta_G)_p, \quad (16.6.5)$$

where Θ_G is the Maurer-Cartan form on G .

2. If Ω^ω is the curvature 2-form of ω , then $\vartheta^*\Omega^\omega$ is the curvature form of $\vartheta^*\omega$ and

$$(\vartheta^*\Omega)_p = \text{Ad}_{\varphi(p)}^{-1} \circ \Omega_p. \quad (16.6.6)$$

Proof. 1. To prove the first equality, we must calculate $\vartheta_*(X)$ for $X \in T_p P$. Let γ be a path in P representing X . Using (16.6.2), we have

$$\begin{aligned} \vartheta_{*p}(X) &= \frac{d}{dt} \Big|_0 \vartheta(\gamma(t)) = \frac{d}{dt} \Big|_0 (\gamma(t) \cdot \varphi(\gamma(t))) = \\ &= \frac{d}{dt} \Big|_0 (\gamma(t) \cdot \varphi(p)) + \frac{d}{dt} \Big|_0 (p, \varphi(\gamma(t))) = \Phi_{\varphi(p)*}(X) + \Phi_{p*} \circ \varphi_{*p}(X). \end{aligned} \quad (16.6.7)$$

The second term describes a vertical vector in $\vartheta(p)$. Thus, we may write it in the form $\Phi_{\vartheta(p)*}(A)$ with $A \in \mathfrak{g}$. Explicitly, since $\varphi_{*p}(X) \in T_{\varphi(p)} G$, we can write

$$\Phi_{p*} \circ \varphi_{p*}(X) = \Phi_{p*} \circ L_{\varphi(p)*} \circ L_{\varphi(p)*}^{-1} \circ \varphi_{*p}(X). \quad (16.6.8)$$

Using

$$L_{\varphi(p)*}^{-1} \circ \varphi_{p*}(Y) = \Theta_G(\varphi_{p*}(X)) = (\varphi^*\Theta_G)_p(X) \quad (16.6.9)$$

and $\Phi_p \circ L_{\varphi(p)} = \Phi_{\vartheta(p)}$, we obtain

$$\vartheta_{p*}(X) = \Phi_{\vartheta(p)*}(X) + \Phi_{\vartheta(p)*}(\varphi^*\Theta_G(X)). \quad (16.6.10)$$

Using this equation, together with the equivariance of ω , we obtain (16.6.5).

2. Using the Structure Equation, we obtain

$$\vartheta^*\Omega = \vartheta^*(d\omega + \frac{1}{2}[\omega, \omega]) = d(\vartheta^*\omega) + \frac{1}{2}[\vartheta^*\omega, \vartheta^*\omega], \quad (16.6.11)$$

that is, $\vartheta^*\Omega$ is the curvature form of $\vartheta\omega$, indeed. To prove (16.6.6), we must calculate $\vartheta^*\Omega(X, Y)$ for any $X, Y \in T_p P$. For that purpose, we use the decomposition (16.6.10) for both vectors. Since Ω is horizontal, only the first terms of this decomposition contribute. Then, using the equivariance of Ω , one immediately obtains (16.6.6). ■

Remark 16.6.2. 1. Sometimes we will use the following short-hand notation for the above transformation laws:

$$\vartheta^*\omega = \text{Ad}_{\varphi}^{-1} \circ \omega + \varphi^*\Theta_G, \quad \vartheta^*\Omega = \text{Ad}_{\varphi}^{-1} \circ \Omega. \quad (16.6.12)$$

In matrix notation, we have $\varphi^*\Theta_G = \varphi^{-1}d\varphi$, and

$$\vartheta^*\omega = \varphi^{-1}\omega\varphi + \varphi^{-1}d\varphi, \quad \vartheta^*\Omega^\omega = \varphi^{-1}\Omega^\omega\varphi. \quad (16.6.13)$$

2. Using the local representative $\rho^\alpha = \varphi \circ s_\alpha$, we read off the following transformation laws for the local representatives of ω and Ω^ω :

$$\mathcal{A}' = \text{Ad}_\rho^{-1} \circ \mathcal{A} + \rho^*\Theta_G, \quad \mathcal{F}' = \text{Ad}_\rho^{-1}\mathcal{F}. \quad (16.6.14)$$

Note that this formula is essentially identical to the transformation law for local representatives from Corollary 16.3.14. We might call the transformation corresponding to a change of local coordinates a *passive gauge transformation*, and a transformation stemming from an actual automorphism of the bundle an *active gauge transformation*.

Since we would like to study bundles up to isomorphism, the action of the gauge group $\text{Gau}(P) = \text{Aut}_M(P)$ on the space of connections $\mathcal{A}(P)$ will be crucial. We will now characterize the infinitesimal version of this action. The gauge algebra $\text{gau}(P)$ can be identified with sections of the adjoint bundle $\text{Ad}(P)$:

$$\text{gau}(P) \cong \Omega^0(M; \text{Ad}(P)) = \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}. \quad (16.6.15)$$

Indeed, each G -invariant vector field $X \in \Gamma(\mathcal{V}P)$, by composition with the global trivialization $\psi^{-1} : \mathcal{V}P \rightarrow P \times \mathfrak{g}$ (cf. Lemma 16.3.1), produces an element of $\Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}$.

An automorphism $\vartheta \in \text{Aut}(P)$ acts on $\mathcal{A}(P)$ by pullbacks along the inverse, $\omega \mapsto \vartheta_*\omega$. For the action of the subgroup $\text{Gau}(P)$, there is an alternative formula for this action which we now derive.

Recall that each vertical automorphism $\vartheta \in \text{Gau}(P)$ corresponds to a unique equivariant function $\varphi : P \rightarrow G$ defined by

$$\varphi \in \text{Hom}_G(P, G), \quad \vartheta(p) = p \cdot \varphi(p)^{-1}. \quad (16.6.16)$$

Here, G acts on itself by conjugations, and the equivariance property reads $\varphi(p \cdot g) = g^{-1}\varphi(p)g$ (which immediately follows from $\vartheta(p \cdot g) = \vartheta(p) \cdot g$). Similarly, each invariant vertical vector field $Y \in \text{gau}(P)$ can be identified, via the global trivialization ψ^{-1} , with a function

$$\tilde{\zeta} := \psi^{-1} \circ Y \in \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}. \quad (16.6.17)$$

Indeed, it is easy to check that $\Phi_g^*\zeta = \text{Ad}_g^{-1}\zeta$. This, in turn, is identified via Proposition 16.2.2 with a section of the associated adjoint bundle:

$$\zeta \in \Omega^0(M; \text{Ad}(P)) = \Gamma^\infty(\text{Ad}(P)). \quad (16.6.18)$$

Explicitly, the identification is $\zeta(m) = \iota_p(\tilde{\zeta}(p)) = [(p, \tilde{\zeta}(p))] \in P \times_G \mathfrak{g}$, where $p \in P_m$. We can write these identifications as isomorphisms

$$\text{gau}(P) \cong \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}} \cong \Gamma^\infty(\text{Ad}(P)). \quad (16.6.19)$$

The next proposition computes the action of the gauge group on the Lie algebra of all infinitesimal automorphisms.

Proposition 16.6.3. *Let $\vartheta \in \text{Gau}(P)$ and $\varphi \in \text{Hom}_G(P, G)$ the corresponding equivariant function. For all $X \in \text{aut}(P) = \mathfrak{X}(P)^G$,*

$$\vartheta_* X - X \in \text{gau}(P). \quad (16.6.20)$$

The corresponding section $\zeta \in \Gamma^\infty(\text{Ad}(P))$ is represented by

$$\tilde{\zeta}(m) = \text{Ad}_{\varphi(m)} \circ (\varphi^* \Theta_G)(X(m)). \quad (16.6.21)$$

Proof. The vector field $\vartheta_* X$ is invariant since ϑ commutes with the principal G -action, and under π_* projects onto the same vector field as X since $\pi \circ \vartheta = \pi$. This shows that $Y = \vartheta_* X - X$ is vertical and G -invariant, thus an element of $\text{gau}(P)$. Now we evaluate both sides of (16.6.5) on the vector field X :

$$\omega(\vartheta_* X) = \text{Ad}_\varphi^{-1} \circ \omega(X) + (\varphi^* \Theta_G)(X), \quad (16.6.22)$$

which is equivalent to

$$\text{Ad}_\varphi \omega(\vartheta_* X) - \omega(X) = \text{Ad}_\varphi \circ (\varphi^* \Theta_G)(X), \quad (16.6.23)$$

By the equivariance of principal connections, $\text{Ad}_\varphi \omega(\vartheta_* X) = \omega(\Phi_{\vartheta_*}^{-1} \vartheta_* X)$, and because ϑ is equivariant and X is invariant, this is just $\omega(\vartheta_* X)$. Therefore

$$\omega(Y) = \text{Ad}_\varphi \circ (\varphi^* \Theta_G)(X). \quad (16.6.24)$$

The connection form was defined as $\omega = \psi^{-1} \circ \mathcal{V}$. Since Y is vertical, $\omega(Y) = \psi^{-1}(Y)$. We get the corresponding Ad-equivariant function

$$\tilde{\zeta} = \psi^{-1} \circ Y = \text{Ad}_\varphi \circ (\varphi^* \Theta_G)(X). \quad (16.6.25)$$

■

Remark 16.6.4. Note that $\text{Ad}_\varphi \circ \varphi^* \Theta_G$ can also be written as $\varphi^* \Theta_G^R$, where Θ_G^R is the right-invariant Maurer-Cartan form on G . Therefore, in matrix notation we may write

$$\tilde{\zeta} = (\mathcal{L}_X \varphi) \cdot \varphi^{-1}.$$

Proposition 16.6.5. *The natural action of $\text{Gau}(P)$ on $\mathcal{A}(P)$ is given by*

$$\vartheta \cdot \omega := (\vartheta^{-1})^* \omega = \text{Ad}_{\varphi} \circ (\omega - \varphi^* \Theta_G), \quad (16.6.26)$$

where $\varphi \in \text{Hom}_G(P, G)$ is the element representing $\vartheta \in \text{Gau}(P)$.

Proof. We will prove the equivalent property $\vartheta^* \omega = \text{Ad}_{\varphi}^{-1} \omega + \varphi^* \Theta_G$. It suffices to check this on invariant vector fields $X \in \mathfrak{X}(G)^G$. By the last proposition, and since $\varphi^* f = \text{Ad}_{\varphi}^{-1} f$ for all $f \in \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}$,

$$(\vartheta^* \omega)(X) = \vartheta^*(\omega(\vartheta_* X)) = \text{Ad}_{\varphi}^{-1}(\omega(\vartheta_* X)) = (\text{Ad}_{\varphi}^{-1} \omega + \varphi^* \Theta_G)(X). \quad (16.6.27)$$

■

We can now characterize the stabilizer subgroups $\text{Gau}(P)_{\omega}$ of gauge transformations that preserve a given connection ω . First consider how gauge transformations affect parallel transport. The following is a simple exercise.

Lemma 16.6.6. *Let $\vartheta \in \text{Gau}(P)$ and $\vartheta_m := \vartheta|_{P_m}$. Then for any path $\gamma : [0, 1] \rightarrow M$,*

$$\text{tra}_{\gamma}^{\vartheta \cdot \omega} = \vartheta_{\gamma(1)} \circ \text{tra}_{\gamma}^{\omega} \circ \vartheta_{\gamma(0)}^{-1}. \quad (16.6.28)$$

Now suppose M is connected and recall the holonomy representation for $m \in M$:

$$\text{Hol}(\omega) : C(m) \rightarrow \text{Diff}(P_m)^G \cong G, \quad \gamma \mapsto \text{tra}_{\gamma}^{\omega}. \quad (16.6.29)$$

Its image $\text{Hol}_p(\omega) < G$ is the holonomy group based at $p \in P_m$, and its identity component $\text{Hol}_p^0(\omega) \triangleleft G$ is a closed subgroup of G . Recall also that $\text{Diff}(P_m)^G \cong G$ is a natural isomorphism, so we will use these two groups interchangeably.

Proposition 16.6.7. *The restriction map $\text{Gau}(P)_{\omega} \rightarrow \text{Diff}(P_m)^G$, $\vartheta \mapsto \vartheta_m := \vartheta|_{P_m}$ is injective and defines an isomorphism of the stabilizer $\text{Gau}(P)_{\omega}$ with the centralizer in G of the holonomy group:*

$$\text{Gau}(P)_{\omega} \cong C_G(\text{Hol}_p(\omega)). \quad (16.6.30)$$

Proof. A gauge transformation preserves ω iff it preserves $\mathcal{H}P$. Equivalently, by Lemma 16.6.6 (applied to ϑ^{-1}), $\vartheta \in \text{Gau}(P)_{\omega}$ iff for all paths $\gamma : [0, 1] \rightarrow M$, ϑ commutes with parallel transport:

$$\vartheta(\gamma(1)) = \text{tra}_{\gamma}^{\omega} \circ \vartheta(\gamma(0)) \circ (\text{tra}_{\gamma}^{\omega})^{-1}. \quad (16.6.31)$$

Taking paths with $\gamma(0) = m$, it follows that if $\vartheta_m = \text{id}_{P_m}$, then $\vartheta = \text{id}_P$. That is, the map $\vartheta \mapsto \vartheta_m$ is injective. Taking $\gamma \in C(m)$, (16.6.31) shows that ϑ_m centralizes $\text{Hol}_p(\omega)$. Conversely, if $\vartheta_m \in C_G(\text{Hol}_p(\omega))$, then we can define ϑ fiberwise by putting

$$\vartheta_{m'} := \text{tra}_\gamma^\omega \circ \vartheta_m \circ (\text{tra}_\gamma^\omega)^{-1}, \quad (16.6.32)$$

where γ is any path $m \rightsquigarrow m'$. The condition $\vartheta \in \text{Gau}(P)_\omega$ guarantees that the right hand side does not depend on the choice of γ and hence globally defines a smooth automorphism. ■

This shows that all stabilizer subgroups are finite dimensional Lie groups. In particular, for an irreducible connection, the stabilizer of a connection ω is isomorphic to the center $Z(G)$. This is another way of detecting reducibility, in addition to the Ambrose-Singer Theorem 16.5.8. Moreover:

Corollary 16.6.8. *The based gauge group $\text{Gau}_m(P) := \{\vartheta \in \text{Gau}(P) \mid \vartheta_m = \text{id}_{P_m}\}$ acts freely on $\mathcal{A}(P)$.*

Corollary 16.6.9. *If G is abelian, every principal connection has stabilizer equal to G , viewed as constant gauge transformations in $\text{Gau}(P) \cong \Gamma^\infty(\text{Ad}(P)) \cong C^\infty(M, G)$.*

Definition 16.6.10 (Moduli space of connections). The (topological) quotient $\mathcal{A}(P)/\text{Gau}(P)$ is called the moduli space of connections.

This moduli space is still “infinite-dimensional”. To get nicer moduli spaces one usually restricts to a sub-class of connections such as flat connections or, more generally, Yang-Mills connections.

16.7 Covariant exterior derivative

In the previous section we have computed the natural action of the gauge group $\text{Gau}(P)$ on the space of connections $\mathcal{A}(P)$. As is clear from Proposition 16.6.5, this is in fact a smooth action of an infinite-dimensional Lie group on an infinite-dimensional smooth manifold. The natural question is then: what is the infinitesimal action, or what are the Killing vector fields corresponding to this action? Since $\mathcal{A}(P)$ is an affine space modeled on $\Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}$, its tangent spaces are naturally identified with the same space:

$$T_\omega \mathcal{A}(P) \cong \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}. \quad (16.7.1)$$

Thus, an infinitesimal gauge transformation $\zeta \in \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}} \cong \text{gau}(P)$ generates a Killing vector field $\zeta_* \in \mathfrak{X}(\mathcal{A}(P)) \cong C^\infty(\mathcal{A}(P), \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}})$. In other words, the map $(\zeta, \omega) \mapsto \zeta_*(\omega)$ is an object that takes a horizontal 0-form ζ (of type Ad) and a connection form ω ,

and produces a horizontal 1-form (again of type Ad). This is exactly the missing analog of an exterior derivative on these spaces of sections! We now turn to studying this object.

Proposition 16.7.1. *The Killing vector field generated by an infinitesimal gauge transformation $\zeta \in \Gamma^\infty(\text{Ad}(P)) \cong \text{gau}(P)$ on $\mathcal{A}(P)$ is given by*

$$-\zeta_*(\omega) = d\tilde{\zeta} + [\omega, \tilde{\zeta}], \quad (16.7.2)$$

where $\tilde{\zeta} \in \Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}$ is the corresponding equivariant function.

Proof. Let $\gamma : (-\epsilon, \epsilon) \rightarrow \text{Gau}(P)$ be a path in $\text{Gau}(P)$ such that $\dot{\gamma}(0) = \zeta$ (where we think of ζ as an element of $\text{gau}(P)$ directly). Let $\tilde{\gamma}$ be the corresponding path in $\Omega_{\text{hor}}^0(P; \mathfrak{g})^{\text{Ad}}$. We have $\dot{\tilde{\gamma}}(0) = \tilde{\zeta}$. Then, by Proposition 16.6.5,

$$-\zeta_*(\omega) = \frac{d}{dt} \Big|_0 (\gamma(-t) \cdot \omega) = \frac{d}{dt} \Big|_0 (\text{Ad}_{\tilde{\gamma}(-t)}(\omega - (\tilde{\gamma}(-t))^* \Theta_G)) = [\omega, \tilde{\zeta}] + d\tilde{\zeta}, \quad (16.7.3)$$

where the last equality follows from the definition of the Lie bracket on a Lie group and the definition of the Maurer-Cartan form. ■

Since, the action of the gauge group is a left action, the minus sign above is convenient to make the map to Killing vector fields a homomorphism of Lie algebras. Now observe that due to $\tilde{\zeta}$ being of type Ad, we can rewrite the Lie bracket above as a Lie derivative along a Killing vector field $\omega(X_p)_*(p) = \psi(\omega(X_p)) = \mathcal{V}(X_p)$ and simplify:

$$(d\tilde{\zeta} + [\omega, \tilde{\zeta}])(X) = d\tilde{\zeta}(X) - \mathcal{L}_{\omega(X)_*} \tilde{\zeta} = d\tilde{\zeta}(X) - d\tilde{\zeta}(\mathcal{V}(X)) = d\tilde{\zeta}(\text{hor } X). \quad (16.7.4)$$

This, as well as the definition of curvature, suggests the following general definition. Note that at this point we have to specialize to associated *vector* bundles.

Definition 16.7.2 (Covariant exterior derivative). Let P be a principal bundle with a principal connection $\mathcal{H}P$ and connection form ω , let V be a finite-dimensional vector space, and let $\alpha \in \Omega^k(P; V)$. The covariant exterior derivative of α with respect to ω is defined as

$$(D^\omega \alpha)(X_0, \dots, X_k) := d\alpha(\text{hor } X_0, \dots, \text{hor } X_k), \quad X_0, \dots, X_k \in \mathfrak{X}(P). \quad (16.7.5)$$

By definition, D^ω fulfills the same product rule (with respect to \wedge) as the ordinary exterior derivative and $D^\omega \alpha$ is horizontal. Moreover, as we will show below, D^ω preserves the symmetry type of any horizontal form.

In particular,

$$\Omega^\omega = D^\omega \omega. \quad (16.7.6)$$

Now we derive a general formula for the covariant exterior derivative.

Proposition 16.7.3 ([17, Prop. 1.4.3]). *Let $P(M, G)$ be a principal bundle, let $G \xrightarrow{\sigma} V$ be a finite-dimensional linear representation, and let ω be a principal connection form on P .*

1. $D^\omega(\Omega_{\text{hor}}^k(P; V)^\sigma) \subset \Omega_{\text{hor}}^{k+1}(P; V)^\sigma$.
2. Let $\tilde{\alpha} \in \Omega_{\text{hor}}^k(P; V)^\sigma$. Then

$$D^\omega \tilde{\alpha} = d\tilde{\alpha} + \sigma_*(\omega) \wedge \tilde{\alpha}, \quad (16.7.7)$$

where

$$(\sigma_*(\omega) \wedge \tilde{\alpha})_p(X_0, \dots, X_k) := \sum_{i=0}^k (-1)^i \sigma_*(\omega_p(X_i)) (\tilde{\alpha}_p(X_0, \dots, \hat{X}_i, \dots, X_k)), \quad (16.7.8)$$

with $p \in P$, $X_0, \dots, X_k \in T_p P$, and where the hat represents an omitted argument.

Proof. 1. Let $\tilde{\alpha} \in \Omega_{\text{hor}}^k(P, V)^\sigma$. By definition of the covariant exterior derivative, $D^\omega \tilde{\alpha}$ is a V -valued horizontal $(k+1)$ -form. We now check equivariance:

$$(\Phi_g^* D^\omega \tilde{\alpha})(X_0, \dots, X_k) = D^\omega \tilde{\alpha}(\Phi_{g*} X_0, \dots, \Phi_{g*} X_k) = \quad (16.7.9)$$

$$= d\tilde{\alpha}(\text{hor } \Phi_{g*} X_0, \dots, \text{hor } \Phi_{g*} X_k) = \quad (16.7.10)$$

$$= d\tilde{\alpha}(\Phi_{g*} \text{hor } X_0, \dots, \Phi_{g*} \text{hor } X_k) = \quad (16.7.11)$$

$$= d(\Phi_g^* \tilde{\alpha})(\text{hor } X_0, \dots, X_k) = \quad (16.7.12)$$

$$= d(\sigma_{g^{-1}} \circ \tilde{\alpha})(\text{hor } X_0, \dots, \text{hor } X_k) = \quad (16.7.13)$$

$$= \sigma_{g^{-1}} \circ D^\omega \tilde{\alpha}(X_0, \dots, X_k). \quad (16.7.14)$$

This shows that $D^\omega \tilde{\alpha}$ is of type σ .

2. Since each of the vectors $X_0, \dots, X_k \in T_p P$ may be decomposed into a vertical and a horizontal part, it is enough to consider the usual cases:

- (a) Let all X_i be horizontal. Let $\omega(X_i) = 0$ and formula (16.7.7) follows from the definition of D^ω .
- (b) Let one of the vectors X_i , say X_0 , be vertical and the rest horizontal. Then, there exists an element $A \in \mathfrak{g}$ such that $X_0 = (A_*)_p = \Phi_*^0(A)$ and a family of vector fields $Y_1, \dots, Y_k \in \mathfrak{X}(M)$ such that their horizontal lifts Y_i^h at p coincide with the vectors X_1, \dots, X_k . Then

$$(D^\omega \tilde{\alpha})(X_0, \dots, X_k) = 0, \quad (\sigma_*(\omega) \wedge \tilde{\alpha})(X_0, \dots, X_k) = \sigma_*(A)(\tilde{\alpha}(X_1, \dots, X_k)). \quad (16.7.15)$$

Using Proposition 10.2.10, Lemma 16.3.10, the horizontality of $\tilde{\alpha}$ and the G -invariance of horizontal lifts, we calculate

$$(d\tilde{\alpha})_p(X_0, \dots, X_k) = (A_*)_p(\tilde{\alpha}(Y_1^h, \dots, Y_k^h)) = \quad (16.7.16)$$

$$= \frac{d}{dt} \Big|_0 \tilde{\alpha}_{p \cdot e^{tA}}(Y_1^h, \dots, Y_k^h) = \quad (16.7.17)$$

$$= \frac{d}{dt} \Big|_0 \tilde{\alpha}_{p \cdot e^{tA}}((\Phi_{e^{tA}})_* X_1^h, \dots, (\Phi_{e^{tA}})_* X_k^h) = \quad (16.7.18)$$

$$= \frac{d}{dt} \Big|_0 \left(\Phi_{e^{tA}}^* \tilde{\alpha} \right)_p (X_1, \dots, X_k) = \quad (16.7.19)$$

$$= \frac{d}{dt} \Big|_0 \sigma_{e^{-tA}} \tilde{\alpha}_p (X_1, \dots, X_k) = \quad (16.7.20)$$

$$= -\sigma_*(A)(\tilde{\alpha}_p(X_1, \dots, X_k)). \quad (16.7.21)$$

Thus, in this case, the right hand side of (16.7.7) also vanishes.

(c) Let at least two vectors X_i be vertical and the rest horizontal. Then

$$D^\omega \tilde{\alpha}(X_0, \dots, X_k) = 0, \quad (\sigma_*(\omega) \wedge \tilde{\alpha})(X_0, \dots, X_k) = 0, \quad (16.7.22)$$

and it remains to show that $d\tilde{\alpha}(X_0, \dots, X_k) = 0$. Since the commutator of vertical vector fields is vertical, the assertion follows from Proposition 10.2.10 and the horizontality of $\tilde{\alpha}$. ■

Definition 16.7.4 (Parallel horizontal form). Let $P(M, G)$ be a principal bundle with a principal connection ω , let $G \xrightarrow{\sigma} V$ and let $E = P \times_G V$ be the associated **VB**. An element $\tilde{\alpha} \in \Omega_{\text{hor}}^k(P, V)^\sigma$ is called parallel with respect to ω if $D^\omega \tilde{\alpha} = 0$.

Remark 16.7.5. 1. In particular, for $k = 0$, this defines the notion of a parallel section of E via the natural isomorphism $\Omega_{\text{hor}}^0(P; V)^\sigma \cong \Gamma^\infty(E)$. Since for $\tilde{s} \in \Omega_{\text{hor}}^0(P; V)^\sigma$, we have $(D^\omega \tilde{s})(X) = d\tilde{s}(\text{hor } X)$ Corollary 16.5.16 implies that this agrees with Definition 16.5.12.
 2. Since the elements of P are interpreted as G -frames on fibers of associated bundles, we may think of parallel local sections of P as local frame fields that are “constant to first order”. Correspondingly, parallel sections of E are then sections whose “coordinates” in such local frame fields have vanishing derivatives. Moreover, by Proposition 16.5.13, specifying local parallel sections is equivalent to specifying the connection.

Proposition 16.7.6 (Bianchi identity [17, Prop. 1.4.11]). *Let P be a principal bundle and let Ω^ω be the curvature form of a principal connection form ω on P . Then Ω^ω is parallel with respect to ω :*

$$D^\omega \Omega^\omega = 0. \quad (16.7.23)$$

Proof. It suffices to show that $d\Omega^\omega(X, Y, Z) = 0$ for horizontal vector fields X, Y, Z on P . Using the Structure Equation, we calculate

$$d\Omega^\omega(X, Y, Z) = \frac{1}{2}d([\omega, \omega])(X, Y, Z) = \quad (16.7.24)$$

$$= X([\omega(Y), \omega(Z)]) - Y([\omega(X), \omega(Z)]) + Z([\omega(X), \omega(Y)]) - \quad (16.7.25)$$

$$- [\omega([X, Y]), \omega(Z)] + [\omega([X, Z]), \omega(Y)] - [\omega([Y, Z]), \omega(X)] = \quad (16.7.26)$$

$$= 0, \quad (16.7.27)$$

because ω vanishes on horizontal vector fields. ■

Remark 16.7.7. As a special case, for $\tilde{\alpha} \in \Omega_{\text{hor}}^k(P; \mathfrak{g})^{\text{Ad}}$, since $\text{ad}(\omega) \wedge \tilde{\alpha} = [\omega, \tilde{\alpha}]$, Proposition 16.7.3 implies the formula we obtained at the beginning of this section:

$$D^\omega \tilde{\alpha} = d\tilde{\alpha} + [\omega, \tilde{\alpha}]. \quad (16.7.28)$$

Thus, in particular, $D^\omega \Omega^\omega = d\Omega^\omega + [\omega, \Omega^\omega]$, and the Bianchi identity takes the form

$$d\Omega^\omega + [\omega, \Omega^\omega] = 0. \quad (16.7.29)$$

Remark 16.7.8. In particular, the covariant exterior derivative of an equivariant map $\tilde{s} \in \text{Hom}_G(P, V)$ is given by

$$D^\omega \tilde{s} = d\tilde{s} + \sigma_*(\omega) \circ \tilde{s}. \quad (16.7.30)$$

Clearly, this is an immediate consequence of formula (16.7.7). It can also be obtained directly in a way completely analogous to (16.7.4).

Applying D^ω to (16.7.7), one finds the following.

Proposition 16.7.9 ([17, Prop. 1.4.13]). *$(D^\omega)^2 = \Omega^\omega$ in the sense that, for $\tilde{\alpha} \in \Omega_{\text{hor}}^k(P; V)$,*

$$D^\omega \circ D^\omega \tilde{\alpha} = \sigma_*(\Omega^\omega) \wedge \tilde{\alpha}. \quad (16.7.31)$$

In particular, the connection is flat iff $(D^\omega)^2 = 0$.

Now let us provide local descriptions of these objects. For simplicity we will omit the index ω . As usual, the local representatives are defined as

$$(D\tilde{\alpha})_\alpha = s_\alpha^* D\tilde{\alpha}. \quad (16.7.32)$$

Let us calculate the right hand side for a 0-form φ explicitly. Formula (16.7.30) implies

$$s_\alpha^*(D\varphi) = d(s_\alpha^*\varphi) + s_\alpha^*(\sigma_*(\omega) \circ \varphi). \quad (16.7.33)$$

For the second term, we calculate

$$(s_\alpha^*\sigma_*(\omega)\varphi)_m(X) = (\sigma_*(\omega)\varphi)_{s_\alpha(m)}(s_{\alpha*}X) = \quad (16.7.34)$$

$$= \sigma_*(\omega_{s_\alpha(m)}(s_{\alpha*}X)\varphi(s_\alpha(m))) = \quad (16.7.35)$$

$$= \sigma_*((s_\alpha^*\omega)_m(X))(s_\alpha^*\varphi)(m) = \quad (16.7.36)$$

$$= \sigma_*(\mathcal{A}_{\alpha,m}(X))\hat{\varphi}_\alpha(m), \quad (16.7.37)$$

where $\mathcal{A}_\alpha = s_\alpha^*\omega$ is the local representative of ω and $X \in T_m M$. Thus, denoting $(D\varphi)_\alpha = D\hat{\varphi}_\alpha$, we have

$$D\hat{\varphi}_\alpha = d\hat{\varphi}_\alpha + \sigma_*(\mathcal{A}_\alpha)\hat{\varphi}_\alpha. \quad (16.7.38)$$

Here, $\sigma_*(\mathcal{A}_*)$ is a 1-form on U with values in $\text{End}(V)$. In the following remark, we analyze formula (16.7.38) further.

Remark 16.7.10. If (U_α, x) is a local chart, $\{\mathbf{t}_a\}$ a basis for \mathfrak{g} and $\{\mathbf{e}_i\}$ a basis in V , we can decompose

$$\hat{\varphi}(x) = \varphi^i(x)\mathbf{e}_i, \quad \mathcal{A}_\alpha = \mathcal{A}_\mu^a dx^\mu \otimes \mathbf{t}_a, \quad D\hat{\varphi}_\alpha = D_\mu \varphi^i dx^\mu \otimes \mathbf{e}_i. \quad (16.7.39)$$

To determine the coefficient functions $D_\mu \varphi^i$, we compute

$$\sigma_*(\mathcal{A}_\alpha)\hat{\varphi}_\alpha = \sigma_*(\mathcal{A}_\mu^a dx^\mu \otimes \mathbf{t}_a)\varphi^i \mathbf{e}_i = \quad (16.7.40)$$

$$= (\mathcal{A}_\mu^a \varphi^i dx^\mu) \otimes (\sigma_*(\mathbf{t}_a)\mathbf{e}_i) = \quad (16.7.41)$$

$$= \mathcal{A}_\mu^a \varphi^i \sigma_{ai}{}^j dx^\mu \otimes \mathbf{e}_j, \quad (16.7.42)$$

with $\sigma_{ai}{}^j$ representing the endomorphism $\sigma_*(\mathbf{t}_a)$ in the basis $\{\mathbf{e}_a\}$. Using this and denoting $\mathcal{A}_{\mu j}^i = \sigma_{aj}{}^i \mathcal{A}_\mu^a$, we obtain

$$D_\mu \varphi^i = \partial_\mu \varphi^i + \mathcal{A}_{\mu,j}^i \varphi^j. \quad (16.7.43)$$

Remark 16.7.11. The local curvature form, expanded in the local basis as in 16.3.12, satisfies

the Structure Equation written as

$$\mathcal{F}_\alpha = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + c_{bc}^a \mathcal{A}_\mu^b \mathcal{A}_\nu^c, \quad (16.7.44)$$

where c_{bc}^a are the structure constants of \mathfrak{g} with respect to the basis $\{\mathbf{t}_a\}$.

Recall that reductions of P to H are in bijective correspondence with smooth sections of the associated bundle $P/H = P \times_G G/H$, cf. Corollary 14.4.6. Since orbits of Lie group actions are weakly embedded submanifolds, we can carry over Proposition 14.4.3 to the case of a general Lie group action $G \curvearrowright F$ by applying it to elements of $\text{Hom}_G(P, F)$ with values in a single orbit of σ .

Proposition 16.7.12 ([17, Prop. 1.6.10]). *Let $P(M, G)$ be a principal bundle and let $G \xrightarrow{\sigma} V$ be a representation. Let $\varphi \in \text{Hom}_G(P, V)$ and assume it takes values in a single orbit O of σ . Let $Q(M, H)$ be the reduction of P defined by φ and some element $f \in O$. Then, a connection $\mathcal{H}P$ is reducible to a connection $\mathcal{H}Q$ iff φ is parallel with respect to $\mathcal{H}P$.*

Proof. Let ω be the connection form for $\mathcal{H}P$ and let (ϑ, i_H) be the morphism corresponding to the reduction Q . Since $Q = \{p \in P : \varphi(p) = f\}$, we have $\vartheta^*\varphi = f$ on Q . Thus,

$$\vartheta^*(D^\omega \varphi) = d(\vartheta^*\varphi) + \sigma_*(\vartheta^*\omega)(\vartheta^*\varphi) = \sigma_*(\vartheta^*\omega)(\vartheta^*\varphi). \quad (16.7.45)$$

Now, $\mathcal{H}P$ is reducible iff $\vartheta^*\omega$ takes values in the Lie algebra \mathfrak{h} of H , that is, iff $\sigma_*(\vartheta^*\omega)(\vartheta^*\varphi) = 0$, that is, iff $\vartheta^*(D^\omega \varphi) = 0$. By the G -equivariance of $\mathcal{H}P$ the latter is equivalent to $D^\omega \varphi = 0$ on the whole of P . ■

Recalling Definition 16.4.8, we conclude that a connection $\mathcal{H}P$ is compatible with $\varphi \in \text{Hom}_G(P, V)$ iff

$$D^\omega \varphi = 0. \quad (16.7.46)$$

16.8 Koszul (linear) connections

Now we show how the covariant exterior derivative induces certain differential operators on the spaces of sections of associated *vector* bundles. As above, let $P(M, G)$ be a principal bundle, $G \xrightarrow{\sigma} V$ be a Lie group representation and let $E = P \times_G V$ be the associated **VB**. Connections on vector bundles induced from principal connections via linear representations are commonly called *linear connections*. The corresponding calculus that we develop below is also known as *Koszul calculus*.

Recall that a principal connection $\mathcal{H}P$ induces a connection $\mathcal{H}E$ and the connection form ω on P induces a connection form $\mathcal{V}^E = \omega^E \in \Omega^1(E; E)$. Using the isomorphism between

$\Omega_{\text{hor}}^k(P; V)^\sigma$ and $\Omega^l(M; E)$ provided by Proposition 16.2.2, we can carry over the notion of covariant exterior derivative to $\Omega^k(M; E)$.

Definition 16.8.1. Let $\alpha \in \Omega^k(M; E)$. The covariant exterior derivative $d^\omega \alpha$ is defined to be the image of $D^\omega \tilde{\alpha}$ under the isomorphism $\Omega_{\text{hor}}^{k+1}(P; V)^\sigma \rightarrow \Omega^{k+1}(M; E)$, that is,

$$\widetilde{d^\omega \alpha} := D^\omega \tilde{\alpha}. \quad (16.8.1)$$

By definition, for $p \in P_m$ and $X_i \in T_m M$, $Y_i \in T_p P$ fulfilling $\pi_*(Y_i) = X_i$, we have

$$(d^\omega \alpha)_m(X_1, \dots, X_{k+1}) = \iota_p \circ (D^\omega \tilde{\alpha})_p(Y_1, \dots, Y_{k+1}). \quad (16.8.2)$$

Since $\Omega^0(M; E) = \Gamma^\infty(E)$ and $\Omega^1(M; E) = \Gamma^\infty(T^* M \otimes E)$, d^ω restricted to 0-forms yields a linear operator from $\Gamma^\infty(E)$ to $\Gamma^\infty(T^* M \otimes E)$.

Definition 16.8.2 (Covariant derivative). The linear operator

$$\nabla^\omega := d^\omega|_{\Omega^0(M; E)} : \Gamma^\infty(E) \rightarrow \Omega^1(M; E) \cong \Gamma^\infty(T^* M \otimes E) \quad (16.8.3)$$

is called the covariant derivative on E induced from ω (not to be confused with the absolute derivative acting on sections of P , cf. Remark 16.5.14).

Exercise 16.5. Verify the following analog of Proposition 10.2.10 expressing the covariant exterior derivative in terms of the covariant derivative and the Lie derivative:

$$\begin{aligned} d^\omega \alpha(X_0, \dots, X_k) &= \sum_{i=1}^k (-1)^i \nabla_{X_i}(\alpha(X_0, \dots, \hat{X}_i, \dots, X_k)) + \\ &\quad + \sum_{i < j} (-1)^{i+j} \alpha([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_k). \end{aligned} \quad (16.8.4)$$

By (16.8.2) and the definition of D^ω , for any $m \in M$ and any $s \in \Gamma^\infty(E)$, we have

$$(\nabla^\omega s)_m(X) = \iota_p \circ (D^\omega \tilde{s})_m(Y) = \iota_p \circ (X_p^h(\tilde{s})), \quad p \in P_m, \quad (16.8.5)$$

where $Y \in T_p P$ fulfilling $\pi_*(Y) = X$ and X^h is the horizontal lift of X to P . In the sequel, we assume that a connection has been chosen and, for simplicity, write ∇ instead of ∇^ω .

Formula (16.8.5) implies a useful expression for the action of ∇ on the local frames of E . To derive it, recall that a local trivialization of P induces one for E . Correspondingly, for a chosen basis $\{\mathbf{e}_i\}$ of the typical fiber V , a local section s_α of P induces a local frame $\{e_i\}_{i=1}^p$ of E via

$$e_i(m) = \iota_{s_\alpha(m)}(\mathbf{e}_i). \quad (16.8.6)$$

Let $\tilde{e}_i : P \rightarrow V$ be the equivariant mapping corresponding to e_i . Then,

$$\tilde{e}_i(s_\alpha(m)) = \mathbf{e}_i. \quad (16.8.7)$$

Proposition 16.8.3 ([17, Prop. 1.5.3]). *In the above setting,*

$$\nabla e_i = \mathcal{A}^j{}_i e_j, \quad (16.8.8)$$

where $\mathcal{A}_\alpha = s_\alpha^* \omega$ is the local representative of ω and $\mathcal{A}^j{}_i$ denotes its matrix with respect to the basis $\{\mathbf{e}_i\}$ of V , cf. Remark 16.7.10.

Proof. Consider (16.8.5) for a point $m \in M$ belonging to the domain of s_α . Since we can take its right hand side at any point in the fiber over m , we calculate it at $s_\alpha(m)$ and for Y we take the vector $s_{\alpha*}(X)$ tangent to the section s_α . Using (16.7.30), (16.8.6) and (16.8.7), for any $X \in T_m M$, we calculate, with \tilde{e}_i denoting the corresponding local equivariant functions on P ,

$$(\nabla e_i)_m(X) = \iota_{s_\alpha(m)}((D\tilde{e}_i)_{s_\alpha(m)}(s_{\alpha*}(X))) = \quad (16.8.9)$$

$$= \iota_{s_\alpha(m)}((d\tilde{e}_i)(s_{\alpha*}(X)) + \sigma_*(\omega(s_{\alpha*}(X)))\tilde{e}_i(s_\alpha(m))) = \quad (16.8.10)$$

$$= \iota_{s_\alpha(m)}(d(s_\alpha^*\tilde{e}_i)(X) + \sigma_*(\mathcal{A}(X))\mathbf{e}_i) = \quad (16.8.11)$$

$$= \iota_{s_\alpha(m)}(\mathcal{A}(X)_i^j \mathbf{e}_j) = \quad (16.8.12)$$

$$= (\mathcal{A}(X)_i^j e_j)(m). \quad (16.8.13)$$

Proposition 16.8.4 ([17, Prop. 1.5.4]). *For any $f \in C^\infty(M)$ and $s \in \Gamma^\infty(E)$, the Leibniz rule holds in the form*

$$\nabla(fs) = df \otimes s + f\nabla s. \quad (16.8.14)$$

Proof. Using Remark 16.2.3, for $m \in M$, $X \in T_m M$ and $p \in P_m$, we calculate

$$(\nabla(fs))_m(X) = \iota_p \circ d(\tilde{f}s)_p(X^h) = \quad (16.8.15)$$

$$= \iota_p \circ d(\tilde{f}\tilde{s})_p(X^h) = \quad (16.8.16)$$

$$= \iota_p \circ ((d\tilde{f})\tilde{s} + \tilde{s}(d\tilde{s}))_p(X^h) = \quad (16.8.17)$$

$$= (df)_m(X)s(m) + f(m)(\nabla s)_m(X). \quad (16.8.18)$$

Remark 16.8.5. 1. Combining Propositions 16.8.3 and 16.8.4 with Remark 16.7.10, for a local section $s = s^i e_i$ of E , decomposed with respect to a local frame e_i , we obtain

$$\nabla s = ds^i \otimes e_i + \mathcal{A}^j{}_i s^i e_j. \quad (16.8.19)$$

2. We have the following obvious generalization of Proposition 16.8.4. For $\alpha \in \Omega^k(M; E)$ and $\beta \in \Omega^l(M)$,

$$d^\omega(\beta \wedge \alpha) = d\beta \wedge \alpha + (-1)^l \beta \wedge d^\omega \alpha. \quad (16.8.20)$$

3. Let E be a \mathbb{K} -vector bundle of rank k over M . Then E is associated to the frame bundle $\text{Fr}(E)$ and by definition, a connection on E is a \mathbb{K} -linear map $\nabla : \Gamma^\infty(E) \rightarrow \Gamma^\infty(T^*M \otimes E)$ fulfilling the Leibniz rule (16.8.14). Then, by the above correspondence, connections on E are in one-to-one correspondence with principal connections on $\text{Fr}(E)$. Thereby, the connection ∇ corresponding to the connection form ω coincides with the covariant derivative defined by ω . Thus, the theory of connections on arbitrary vector bundles boils down to the theory of covariant derivatives on associated vector bundles.
4. By point 3, Proposition 16.8.3 immediately extends to any **VB** E endowed with a connection. Then, P coincides with $\text{Fr}(E)$ and σ is the defining representation of $\text{GL}_n(\mathbb{K})$.

The following proposition clarifies the relation between the covariant derivative and the connection form ω^E , see (16.4.7).

Proposition 16.8.6 ([17, Prop. 1.5.6]). *For $X \in \mathfrak{X}(M)$,*

$$\nabla s(X) = \omega^E(s_*(X)). \quad (16.8.21)$$

Proof. Combining Proposition 16.5.13, (16.4.7) and (16.8.5), we get

$$\omega^E(s_*(X)) = \iota_P \circ (\iota_P)_*^{-1}(\mathcal{V}^E(s_*(X))) = \iota_P \circ X^h(\tilde{s}) = \nabla s(X). \quad (16.8.22)$$

Remark 16.8.7. This implies that Proposition 16.5.13 could be viewed as the definition of a covariant derivative of sections of arbitrary **FBs**, not just **VBS**. However, such an object would take values in $\mathcal{V}E$ instead of E itself, precluding the existence of any Koszul-like calculus.

Next recall the notion of parallelity. By (16.8.5), a section $s \in \Gamma^\infty(E)$ is parallel iff $\nabla_X s = 0$ for all $X \in \mathfrak{X}(M)$. Proposition 16.8.6 implies the following linear version of Corollary 16.5.16.

Corollary 16.8.8 ([17, Cor. 1.5.7]). A section $s \in \Gamma^\infty(E)$ is parallel with respect to a connection $\mathcal{H}E$ in the sense of Definition 16.7.4 ($D^\omega \tilde{s} = 0$) iff it is parallel in the sense of Definition 16.5.12 ($\text{im}(s_{*m}) \subset \mathcal{H}_{s(m)} E$ for all $m \in M$).

It will often be useful to view the covariant derivative as a differential operator acting on sections: for every $X \in \mathfrak{X}(M)$, the covariant derivative induces a map

$$\nabla_X : \Gamma^\infty(E) \rightarrow \Gamma^\infty(E), \quad \nabla_X s := \nabla s(X). \quad (16.8.23)$$

Remark 16.8.9. Since describing parallel sections is equivalent to specifying the horizontal distribution, the operator ∇ itself is often called a *linear connection*.

Proposition 16.8.10 ([17, Prop. 1.5.8]). For $X, X_1, X_2 \in \mathfrak{X}(M)$, $s, s_1, s_2 \in \Gamma^\infty(E)$ and $f \in C^\infty(M)$,

1. $\nabla_{X_1+X_2}s = \nabla_{X_1}s + \nabla_{X_2}s$,
2. $\nabla_X(s_1 + s_2) = \nabla_Xs_1 + \nabla_Xs_2$,
3. $\nabla_f Xs = f\nabla_Xs$,
4. $\nabla_X(fs) = f\nabla_Xs + X(f)s$.

Proof. Points 1 and 2 are immediate consequences of the definition of ∇_X . Using Remark 16.2.3, together with $(fX)^h = \tilde{f}X^h$, we get

$$(\nabla s)_m(fX) = \iota_p \circ ((fX)^h(\tilde{s})) = \iota_p \circ (\tilde{f}X^h(\tilde{s})) = f(m)(\nabla s)_m(X), \quad (16.8.24)$$

for any $p \in P_m$. This proves point 3. Point 4 follows from Proposition 16.8.4. ■

Remark 16.8.11. 1. By the locality property of Proposition 16.8.10(3), at any $m \in M$ the value of $(\nabla_Xs)(m)$ depends only on the value of X at m and the values of s along any smooth curve representing X_m . Thus this object is tensorial in X and we obtain a map $\nabla : TM \times \Gamma^\infty(E) \rightarrow E$ defined by

$$\nabla_{Y_m}s = (\nabla_Xs)(m), \quad (16.8.25)$$

where X is an arbitrary extension of the tangent vector $Y_m \in T_m M$ to a smooth vector field on M .

2. The covariant derivative on E naturally induces covariant derivatives on all tensor bundles over E : for the dual bundle E^* we define

$$(\nabla_X^{E^*} s^*)(s) := X(\langle s^*, s \rangle) - \langle s^*, \nabla_X^E s \rangle, \quad (16.8.26)$$

where $X \in \mathfrak{X}(M)$, $s \in \Gamma^\infty(E)$ and $s^* \in \Gamma^\infty(E^*)$. Next, we extend ∇_X to any tensor product built from E and E^* by requiring that it be a derivation with respect to tensor products.

3. Let E_1, E_2 be two **VBs** over M endowed with connections ∇^1 and ∇^2 . Then

$$\nabla(s_1 \otimes s_2) := (\nabla^1 s_1) \otimes s_2 + s_1 \otimes (\nabla^2 s_2), \quad s_i \in \Gamma^\infty(E_i), i = 1, 2 \quad (16.8.27)$$

defines the *tensor product connection* $\nabla = \nabla^1 \otimes \nabla^2$ on $E_1 \otimes E_2$.

In particular, let E_1 and E_2 be associated with principal bundles $P_1(M, G_1)$ and $P_2(M, G_2)$, respectively. Then, by Example 14.2.6(3), $E_1 \otimes E_2$ is naturally associated with the fiber product $P_1 \times_M P_2$, cf. 14.1.10(2). If ω_i are principal connection forms on P_i , then $P_1 \times_M P_2$ is endowed the natural connection form $\vartheta^* \omega$ given by (16.4.17). If $\nabla^i, i = 1, 2$, are the covariant derivatives on E_i induced from ω_i , respectively, then the covariant derivative induced from $\vartheta^* \omega$ coincides with the tensor product connection $\nabla^1 \otimes \nabla^2$ (exercise).

Next recall that, as a consequence of Proposition 16.7.9, the square of the exterior covariant derivative is related to the curvature. Let us find the analog of this in Koszul calculus. Recall that Ω^ω is a horizontal 2-form on P with values in \mathfrak{g} , thus $\sigma_*(\Omega^\omega)$ takes values in $\text{End}(V)$. By Proposition 16.2.2, to Ω^ω there corresponds a 2-form on M with values in the endomorphism bundle $\text{End}(E) = E \otimes E^*$:

$$R_m^\nabla(X, Y) := \iota_p \circ \sigma_*(\Omega^\omega(X^h, Y^h)) \circ \iota_p^{-1}, \quad (16.8.28)$$

where $m \in M$, $p \in P_m$, $X, Y \in T_m M$ and X^h, Y^h are their horizontal lifts to p .

Definition 16.8.12 (Curvature endomorphism). The 2-form R^∇ is called the curvature endomorphism form associated with Ω^ω .

Proposition 16.8.13 ([17, Prop. 1.5.11]). *For any $X, Y \in \mathfrak{X}(M)$,*

$$R^\nabla(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]}. \quad (16.8.29)$$

Proof. Let X^h, Y^h be the horizontal lifts to P and let $p \in P$. Let $\tilde{s} \in \text{Hom}_G(P, V)$ be an equivariant function and $s \in \Gamma^\infty(E)$ the associated section. Using (16.3.2), (16.8.5) and $\text{hor}([X^h, Y^h]) = [X, Y]^h$, we calculate

$$\Phi_*^P(\Omega^\omega(X^h, Y^h))\tilde{s}(p) = -\text{ver}([X^h, Y^h])_p(\tilde{s}) = \quad (16.8.30)$$

$$= -[X^h, Y^h]_p(\tilde{s}) + \text{hor}([X^h, Y^h])_p(\tilde{s}) = \quad (16.8.31)$$

$$= -[X^h, Y^h]_p(\tilde{s}) + [X, Y]^h_p(\tilde{s}) = \quad (16.8.32)$$

$$= -X_p^h(Y^h(\tilde{s})) + Y_p^h(X^h(\tilde{s})) + [X, Y]^h_p(\tilde{s}) = \quad (16.8.33)$$

$$= -\iota_p^{-1} \circ (\nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s)(m). \quad (16.8.34)$$

The assertion follows from $\Phi_*^P(A)(\tilde{s}) = -\sigma_*(A)\tilde{s}(p)$ for all $A \in \mathfrak{g}$. ■

Remark 16.8.14. 1. Viewing the covariant derivative as a linear map

$$\nabla : \mathfrak{X}(M) \rightarrow \text{End}(\Gamma^\infty(E)), \quad X \mapsto \nabla_X, \quad (16.8.35)$$

we conclude that this map is a Lie algebra homomorphism iff the curvature form R^∇ vanishes.

2. Formula (16.8.29) extends to sections in arbitrary tensor bundles $\mathbb{T}_l^k E$ over E , where R^∇ acts on $\mathbb{T}_l^k E$ by the representation induced by σ_* .

Example 16.8.15 (Canonical connections on a Lie group, part 3). We pick up Example 16.5.18. For each λ we get an induced linear connection ∇^λ on $P \times_G \mathfrak{g}$, which we identify with TG by means of the isomorphism ϑ described in Example 14.2.14.

Since left-invariant vector fields form a global frame for TG , it suffices to compute it for left-invariant fields. The derivative of arbitrary vector fields can then be computed by linearity and via the Leibniz rule. Let $A, B \in \mathfrak{g}$ and let $A_L, B_L \in \mathfrak{X}(G)$ be the corresponding left-invariant vector fields. For each $B \in \mathfrak{g}$ we then have a unique equivariant function $\tilde{B}_L : G \times G \rightarrow \mathfrak{g}$ which we identified in Example 14.2.19:

$$\tilde{B}_L(g_1, g_2) = \text{Ad}_{g_2}^{-1} B. \quad (16.8.36)$$

Now we compute the covariant derivative of this function. We will do so in the direction of a tangent vector $(Y_1 + Y_2) \in T_{g_1}G \times T_{g_2}G$:

$$D^\lambda \tilde{B}_L(Y_1 + Y_2) = (\text{d}\tilde{B}_L + \omega_\lambda \tilde{B}_L)(Y_1 + Y_2). \quad (16.8.37)$$

We use the fact that $\text{d}(\text{Ad}_{g_2}^{-1} B)(Y_1 + Y_2) = -[L_{g_2}^{-1} Y_2, \text{Ad}_{g_2}^{-1} B]$ and that ω_λ acts on \tilde{B}_L by adjoint (Lie bracket) to compute

$$D^\lambda \tilde{B}_L(Y) = \frac{1+\lambda}{2} [\theta(Y), \text{Ad}_{g_2}^{-1} B]. \quad (16.8.38)$$

Assuming $\pi_* Y = X \in TG$, we have

$$\begin{aligned} (\nabla_X \tilde{B}_L)_{g_1 g_2^{-1}} &= \iota_{(g_1, g_2)} \circ (D^\lambda \tilde{B}_L)(Y) = \left[\left((g_1, g_2), D^\lambda \tilde{B}_L(Y) \right) \right] = \\ &= \left[\left((g_1 g_2^{-1}, e), \frac{1+\lambda}{2} [\text{Ad}_{g_2} \theta(Y), B] \right) \right], \end{aligned} \quad (16.8.39)$$

where θ is the horizontal 1-form defined in Example 16.2.6. It is most convenient to pick the lift of $X \in T_g G$ to $(g, e) \in P \times_G \mathfrak{g}$ as $Y_1 = X$ and $Y_2 = 0$, so that $\theta(Y) = L_{g_*}^{-1} X = \Theta_G(X)$.

Therefore we find, for $X = A_L$, and using again the isomorphism ϑ ,

$$\nabla_{A_L} B_L = \left[\left((g, e), \frac{1+\lambda}{2} [L_{g^*}^{-1} A_L, B] \right) \right] = \frac{1+\lambda}{2} [A, B]_L. \quad (16.8.40)$$

By an analogous computation, it is also easy to compute the covariant derivative of right-invariant vector fields A_R :

$$\nabla_{A_R} B_R = \frac{\lambda-1}{2} [A, B]_R, \quad (16.8.41)$$

however note that $[A, B]_R = -[A_R, B_R]$. This agrees with our earlier finding that under ∇^- left-invariant vector fields are parallel, and under ∇^+ the right-invariant ones are.

Now it is easy to compute the curvature tensor on invariant vector fields (which, again, is sufficient because they form a global frame) using Proposition 16.8.13:

$$R^\lambda(A_L, B_L)C_L = -\frac{1-\lambda^2}{4} [[A, B], C]_L, \quad R^\lambda(A_R, B_R)C_R = \frac{1-\lambda^2}{4} [[A, B], C]_R. \quad (16.8.42)$$

For arbitrary vectors $X, Y, Z \in T_g G$ this can be written as

$$\begin{aligned} R^\lambda(X, Y)Z &= \\ &= -\frac{1-\lambda^2}{4} L_{g^*} [[L_{g^*}^{-1} X, L_{g^*}^{-1} Y], L_{g^*}^{-1} Z] = \frac{1-\lambda^2}{4} R_{g^*} [[R_{g^*}^{-1} X, R_{g^*}^{-1} Y], R_{g^*}^{-1} Z]. \end{aligned} \quad (16.8.43)$$

We will continue studying these connections in Example 17.1.16.

16.9 Parallel transport

Let us now look in more detail at pullback connections and associated covariant derivatives. Let $P(M, G)$ be a principal bundle, $f : N \rightarrow M$ a smooth map, and f^*P the pullback bundle. Let $G \xrightarrow{\sigma} V$ be a Lie group representation and $E = P \times_G V$ the corresponding associated bundle. By Example 14.2.6(2), the pullback bundle f^*E is naturally associated with f^*P via the **VB** isomorphism

$$f^*E \rightarrow f^*P \times_G V, \quad (y, [(p, f)]) \mapsto [(y, p), f], \quad (16.9.1)$$

cf. (14.2.16). Recall the commuting diagram

$$\begin{array}{ccc} f^*E & \xrightarrow{\pi_2} & E \\ \downarrow \pi_1 & & \downarrow \pi_V \\ N & \xrightarrow{f} & M. \end{array} \quad (16.9.2)$$

By Corollary 16.4.3(2), every principal connection ω on P induces a connection $\vartheta^*\omega$ on f^*P , where $\vartheta : f^*P \rightarrow P$ is the induced bundle morphism. The connection $\mathcal{H}(f^*E)$ induced from $\vartheta^*\omega$ is given by

$$\mathcal{H}(f^*E) = (\pi_2)_*^{-1}(\mathcal{H}E). \quad (16.9.3)$$

Using the obvious identification

$$T_{(y,e)}f^*E \cong (\pi_1)_*(T_{(y,e)}f^*E) \oplus (\pi_2)_*(T_{(y,e)}f^*E), \quad (16.9.4)$$

we obtain

$$T_{(y,e)}f^*E = \{(Y, Z) \in T_yN \oplus T_eE : f_{*y}(Y) = (\pi_V)_{*e}(Z)\}. \quad (16.9.5)$$

Thus the decomposition $(Y, Z) \in T_{(y,e)}f^*E$ with respect to $\mathcal{H}(f^*E)$ is given by

$$(Y, Z) = (0, Z^v) + (Y, Z^h), \quad (16.9.6)$$

with $Z = Z^v + Z^h$ being the decomposition w.r.t. $\mathcal{H}E$. Let us analyze the induced covariant derivative $\nabla^{f^*\omega}$. For that purpose, it is convenient to view the space of sections of f^*E as follows.

Definition 16.9.1 (Section along a map). In the above notation, a section of E along f is a map $\phi : N \rightarrow E$ fulfilling

$$\pi_V \circ \phi = f. \quad (16.9.7)$$

The vector space of sections of E along f is denoted $\Gamma_f^\infty(E)$.

Clearly ϕ is a section of E along f iff $y \mapsto (y, \phi(y))$ is a section of f^*E , that is, $\Gamma_f^\infty(f^*E)$ is canonically isomorphic to $\Gamma_f^\infty(E)$.

Now, let $s \in \Gamma_f^\infty(f^*E)$ and let $Y \in \mathfrak{X}(N)$. Representing s by a section $\phi \in \Gamma_f^\infty(E)$ and using Proposition 16.8.6, together with (14.2.16), (16.9.6) and (16.4.7), we calculate

$$(\nabla^{\vartheta^*\omega}s)_{(y,e)}(Y) = \omega_{(y,e)}^{f^*E}(s_*(Y)) = \quad (16.9.8)$$

$$= \omega_{(y,e)}^{f^*E}((Y, \phi_*(Y))) = \quad (16.9.9)$$

$$= \iota_{(y,p)} \circ (\iota_{(y,p)})_*^{-1}(0, (\phi_*(Y))^v) = \quad (16.9.10)$$

$$= (y, \iota_p \circ (\iota_p)_*^{-1}(\phi_*(Y))^v) = \quad (16.9.11)$$

$$= (y, \omega_e^E(\phi_*(Y))). \quad (16.9.12)$$

We see that, associated with $\nabla_Y^{f^*\omega}$, there is an operator

$$\nabla_Y^f : \Gamma_f^\infty(E) \rightarrow \Gamma_f^\infty(E), \quad \nabla_Y^f \phi := \omega_e^E(\phi_*(Y)). \quad (16.9.13)$$

Definition 16.9.2. The operator ∇^f is called the covariant derivative along the map f .

We have

$$\nabla_Y^{f^*\omega}(\text{id}_N \times \phi) = \text{id}_N \times \nabla_Y^f \phi, \quad (16.9.14)$$

and, by construction, ∇_Y^f inherits the properties listed in Proposition 16.8.10. Moreover, it fulfills an obvious chain rule: for another map $q : L \rightarrow N$, the composition $\phi \circ q$ is a section along $f \circ q$ and for $X \in TL$ we have

$$\nabla_X^{f \circ q}(\phi \circ q) = \nabla_{q_*(X)}^f \phi. \quad (16.9.15)$$

Finally, the following proposition gives a geometric interpretation of the covariant derivative in terms of parallel transport.

Proposition 16.9.3 ([17, Prop. 1.7.17]). *Let ω^E be a connection on E , let ∇ be its covariant derivative and let $s \in \Gamma^\infty(E)$. Then, for any $m \in M$ and any $X \in \mathfrak{X}(M)$,*

$$\nabla_X s(m) = \frac{d}{dt} \Big|_0 \left(\text{tra}_\gamma^{\omega^E}(t) \right)^{-1} \circ s(\gamma(t)), \quad (16.9.16)$$

where $\gamma : I_\epsilon \rightarrow M$ is an integral curve of X through $m = \gamma(0)$ and $I_\epsilon = (-\epsilon, \epsilon)$.

Proof. Let γ^h be the horizontal lift of γ to P starting at $p \in P_m$. Let X^h be the horizontal lift of X to P . Then,

$$X_p^h(\tilde{s}) = \frac{d}{dt} \Big|_0 \tilde{s} \circ \gamma^h(t) = \frac{d}{dt} \Big|_0 \left(\iota_{\gamma^h(t)} \right)^{-1} s(\gamma(t)) \quad (16.9.17)$$

and, thus, (16.8.5), (16.5.18) imply the assertion. ■

Rewriting (16.9.16) as

$$\nabla_X s(m) = \lim_{t \rightarrow 0} \frac{(\text{tra}_\gamma^{\omega^E}(t))^{-1} \circ s(\gamma(t)) - s(\gamma(0))}{t}, \quad (16.9.18)$$

we get a geometric interpretation of the covariant derivative. In particular, note that a section s is parallel iff the curve $s \circ \gamma$ in E is horizontal for any integral curve γ of X .

Now let us consider an arbitrary smooth path $\gamma : I \rightarrow M$. By definition, a section of E along γ is a map $\phi : I \rightarrow E$ such that $\pi_V \circ \phi = \gamma$. We have the canonical isomorphism between

$\Gamma_\gamma^\infty(E)$ and $\Gamma^\infty(\gamma^*E)$, and there is an associated covariant derivative along γ . According to (16.9.13), it is given by

$$\frac{D\omega}{D\gamma(t)} := \nabla_{\partial_t}^\gamma : \Gamma_\gamma^\infty(E) \rightarrow \Gamma_\gamma^\infty(E), \quad \frac{D\omega}{D\gamma(t)}\phi = \omega^E(\phi_*(\partial_t)), \quad (16.9.19)$$

where ∂_t is the standard unit vector field on \mathbb{R} and ω^E is the connection form on E . Now, clearly, for any $s \in \Gamma^\infty(E)$,

$$\phi = s \circ \gamma \quad (16.9.20)$$

is a section of E along γ and, for this choice of ϕ , we get

$$\frac{D\omega}{D\gamma(t)}(s \circ \gamma) = \omega^E(s_*(\dot{\gamma})) = \omega^E(\partial_t(s \circ \gamma)). \quad (16.9.21)$$

Thus, a section $s \circ \gamma$ of E along f is parallel iff

$$\frac{D\omega}{D\gamma(t)}(s \circ \gamma) = 0. \quad (16.9.22)$$

To summarize, we have the following.

Proposition 16.9.4. *The parallel transport operator $\text{tra}_\gamma^{\omega^E} : E_{\gamma(0)} \rightarrow E_{\gamma(1)}$ along a path γ is given by the set of solutions of the differential equation (16.9.22) with the initial condition $s(\gamma(0))$ spanning the fiber $E_{\gamma(0)}$.*

Remark 16.9.5. In particular, if γ is an integral curve of a vector field $X \in \mathfrak{X}(M)$, then $\text{tra}_\gamma^{\omega^E}(t)(e)$, $e \in E_{\gamma(0)}$, is an integral curve of the horizontal lift $X^h \in \mathfrak{X}(E)$. We can write this symbolically for the entire flow F^X of X as

$$\text{tra}_{F^X}^{\omega^E}(t) = F_t^{X^h}. \quad (16.9.23)$$

Remark 16.9.6 (Synchronous framing). For a **VB** $E \rightarrow M$ with connection ∇ , let ω be the principal connection form on the frame bundle $\text{Fr}(E)$ and Ω^ω its curvature. Denote $\dim M = n$ and consider an open ball $B \subset \mathbb{R}^n$ centered at 0. Let x^1, \dots, x^n be the standard coordinates on B and let

$$X^r := \sum_i x^i \partial_i \quad (16.9.24)$$

be the radial vector field on B . Let (U, κ) be a local chart sending U to B . Bia κ , parallel transport along rays $t \mapsto t\mathbf{x}$, $\mathbf{x} \in B$, provides a local trivialization of E over B by identifying the fibers $E_{\mathbf{x}}$ with E_0 . The corresponding local frame is called *synchronous*.

Let $\mathbb{A} = \kappa^*\mathcal{A}$ be the local representative of ω on B with respect to a synchronous frame $\{e_i\}$,

viewed as a local section of $\text{Fr}(E)$, and let $\mathbb{F} = \kappa^*\mathcal{F}$ be the corresponding representative of Ω^ω . Then (16.8.8) implies

$$\nabla_{X'} e_i = \mathbb{A}^j{}_i(X^r) e_j = 0, \quad (16.9.25)$$

that is, $\mathbb{A}(X^r) = 0$. Thus implies

$$\mathcal{L}_{X'} \mathbb{A} = d\mathbb{A}(X^r) = \mathbb{F}(X^r). \quad (16.9.26)$$

We decompose $\mathbb{A} = \mathbb{A}_i dx^i$, $\mathbb{F} = \frac{1}{2}\mathbb{F}_{ij} dx^i \wedge dx^j$. Then

$$\mathcal{L}_{X'} \mathbb{A} = X^r (\mathbb{A}_i) dx^i + \mathbb{A}_i dx^i. \quad (16.9.27)$$

Comparing these two formulas, we read off

$$X^r (\mathbb{A}_i) + \mathbb{A}_i = -\mathbb{F}_{ij} x^j. \quad (16.9.28)$$

This implies

$$\mathbb{A}_i(\mathbf{x}) = -\frac{1}{2}\mathbb{F}_{ij}(0)x^j + O(\|\mathbf{x}\|^2). \quad (16.9.29)$$

In particular, $\mathbb{A}(0) = 0$.

17 Cartan geometry

17.1 Affine connections

Up till now, we have been studying connections on arbitrary bundles, and the properties of connections were essentially unrelated to the base manifold M . In this section we consider connections on TM (and more generally on tensor and jet bundles over M), which will allow us to connect certain properties of such connections to intrinsic structures on M itself.

Since the association functor is an equivalence of categories, connections on arbitrary bundles induce principal connections on corresponding principal bundles (provided the structure group is finite-dimensional and the group action on the typical fiber is at least infinitesimally faithful). This is especially easy to see in the case of VBs. Recall that a local frame on a vector bundle of rank k consists of k linearly independent local sections. This is in fact equivalent to the definition of a local frame as a local section of $\text{Fr}(E)$: every such section induces a local trivialization of E , and picking a basis $\{\mathbf{e}_i\}_{i=1}^k$, the preimages of these basis vectors under this local trivialization are k linearly independent local sections of E . The explicit formula was given in (16.8.7).

Now, since a linear connection on E amounts to a specification of parallel sections, we can define the corresponding principal connection on the frame bundle $\text{Fr}(E)$ by declaring a

local frame (local section of $\text{Fr}(E)$) to be parallel iff it consists of k parallel sections. In particular, linear connections on the tangent bundle TM are in one-to-one correspondence with principal connections on the frame bundle $\text{Fr}(M)$. The latter connections are briefly referred to as *linear connections on M* .

Let M be an n -dimensional manifold and let $\text{Fr } M$ be its frame bundle. Recall that a linear frame at $m \in M$ is an ordered basis $u = (u_1, \dots, u_n)$ in $T_m M$ and that $\pi : \text{Fr } M \rightarrow M$, $\pi(u) = m$, is a principal $\text{GL}_n(\mathbb{R})$ -bundle. The free right action of $\text{GL}_n(\mathbb{R})$ on $\text{Fr } M$ is given by

$$\Phi : \text{Fr } M \times \text{GL}_n(\mathbb{R}) \rightarrow \text{Fr } M, \quad (u, g) \mapsto ug = (u_1g_1^i, \dots, u_ng_n^i). \quad (17.1.1)$$

In the sequel, the defining representation of $\text{GL}_n(\mathbb{R})$ on \mathbb{R}^n will be denoted by σ_n^0 , so $\sigma_n^0(g)\mathbf{x} = g\mathbf{x}$.

Unlike general PFBs, the frame bundle of M is naturally associated with the tangent bundle TM , i.e. there is a natural isomorphism between TM and $\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n$, given by

$$\vartheta^{-1} : \text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n \rightarrow TM, \quad \vartheta^{-1}([(u, \mathbf{x})]) := x^i u_i, \quad (17.1.2)$$

where x^i are the components of $\mathbf{x} \in \mathbb{R}^n$ in the standard basis $\{\mathbf{e}_i\}$ of \mathbb{R}^n . It is easy to see that ϑ^{-1} is an isomorphism, and via dualization and tensor products it produces corresponding isomorphisms for all tensor bundles: $T_l^k M \cong \text{Fr } M \times_{\text{GL}_n(\mathbb{R})} T_l^k \mathbb{R}^n$. The presence of the isomorphism ϑ allows us to identify each frame $u \in \text{Fr } M$ with the isomorphism that computes the coordinates of a given tangent vector in TM in the frame u :

$$\iota_u^\vartheta := \vartheta^{-1} \circ \iota_u : \mathbb{R}^n \rightarrow T_{\pi(u)} M, \quad \mathbf{x} \mapsto \sum_{i=1}^n x^i u_i. \quad (17.1.3)$$

The isomorphism ι_u^ϑ is often identified in literature with the frame u itself. We stick to this more explicit notation for clarity.

The isomorphism ϑ can be viewed as a 1-form on TM with values in this associated bundle, and by Proposition 16.2.2 it corresponds to the following \mathbb{R}^n -valued horizontal 1-form on the frame bundle.

Definition 17.1.1 (Soldering on the frame bundle). The differential form $\theta \in \Omega^1(\text{Fr } M, \mathbb{R}^n)$ defined by

$$\theta(Y) := (\iota_u^\vartheta)^{-1} \circ \pi_*(Y), \quad Y \in T_u \text{Fr } M, \quad (17.1.4)$$

is called the *canonical \mathbb{R}^n -valued 1-form* on $\text{Fr } M$, or the soldering form.

Proposition 17.1.2 ([17, Prop. 2.1.4]). θ belongs to $\Omega_{\text{hor}}^1(\text{Fr } M; \mathbb{R}^n)^{\sigma_n^0}$.

Proof. By definition, θ is horizontal. Let $u \in \text{Fr } M$ and $g \in \text{GL}_n(\mathbb{R})$. If we view u as a map $T_{\pi(u)}M \rightarrow \mathbb{R}^n$, then to $\Phi_g(u) = ug$ there corresponds the map

$$\iota_u^\vartheta \circ g : \mathbb{R}^n \xrightarrow{g} \mathbb{R}^n \xrightarrow{\iota_u^\vartheta} T_{\pi(u)}M. \quad (17.1.5)$$

Thus, for any $Y \in T_u \text{Fr } M$,

$$\begin{aligned} (\Phi_g^* \theta)_u(Y) &= \theta_{ug}(\Phi_{g*} Y) = (\Phi_g(u))^{-1}(\pi_* \circ \Phi_{g*}(Y)) = \\ &= (\iota_u^\vartheta \circ g)^{-1}(\pi_*(Y)) = g^{-1}\theta_u(Y). \end{aligned} \quad (17.1.6)$$

■

Remark 17.1.3. By Proposition 16.2.2, via the isomorphism (17.1.2), to θ there corresponds a unique 1-form $\hat{\theta} \in \Omega^1(M; TM)$ given by

$$\hat{\theta}_m(X) = \iota_u^\vartheta \circ \theta(Y) = \iota_u^\vartheta \circ (\iota_u^\vartheta)^{-1} \circ \pi_*(Y) = X, \quad (17.1.7)$$

where $\pi(u) = m$, $X \in T_m M$ and $Y \in T_u \text{Fr } M$ fulfilling $\pi_*(Y) = X$. Thus, $\hat{\theta}(X) = X$. That is why $\hat{\theta}$ is usually called the *tautological 1-form*. Note that it coincides with the identity tensor, $\hat{\theta} = \text{id}_{TM}$.

Thus, more generally, we can define a soldering as a choice of an isomorphism between TM and a **VB** associated to some **PFB**, and the corresponding horizontal 1-form on the **PFB**. If E is a **VB** associated to a principal bundle P via $E = P^{[\sigma]} = P \times_G V$, then via Proposition 16.2.2, to ϑ there corresponds a unique horizontal 1-form $\theta \in \Omega_{\text{hor}}^1(P; V)^\sigma$. This leads to the following definition of soldering on general principal bundles.

Definition 17.1.4 (Soldering on **PFBs**). Let P be a principal G -bundle, let $G \xrightarrow{\sigma} V$ be a linear representation, and let $E = P^{[\sigma]} = P \times_G V$ be the associated **VB**. Then a soldering form on P is a horizontal 1-form of type σ , $\theta \in \Omega_{\text{hor}}^1(P; V)^\sigma$, such that the corresponding bundle morphism $\vartheta \in \Omega^1(M; E) = \text{Hom}(TM, E)$ is an isomorphism. In particular, $\dim V = \dim M$. It corresponds to a unique $\vartheta \in \Omega^1(M; E)$, which defines an isomorphism of **VBs** $\vartheta : TM \rightarrow E$ and is also called a *soldering*.

Definition 17.1.5 (Torsion form). Let $\mathcal{H}M$ be a linear connection on M and let ω be its principal connection form. If $\theta \in \Omega_{\text{hor}}^1(\text{Fr } M; \mathbb{R}^n)^{\sigma_n^0}$ is the canonical 1-form on $\text{Fr } M$, then the 2-form $\Theta \in \Omega_{\text{hor}}^2(\text{Fr } M, \mathbb{R}^n)^{\sigma_n^0}$ defined by

$$\Theta := D^\omega \theta \quad (17.1.8)$$

is called the torsion form of $\mathcal{H}M$.

The Structure Equation 16.3.19 for the curvature is supplemented by a structure equation involving the torsion form.

Proposition 17.1.6 (Structure Equations [17, Prop. 2.1.11]). *Let ω, Ω and Θ be, respectively, the connection, curvature and torsion forms of a linear connection \mathcal{HM} . Then, for any $X, Y \in T_u \text{Fr } M$,*

$$\Omega(X, Y) = d\omega(X, Y) + [\omega(X), \omega(Y)], \quad (17.1.9)$$

$$\Theta(X, Y) = d\theta(X, Y) + (\omega(X)\theta(Y) - \omega(Y)\theta(X)). \quad (17.1.10)$$

Proof. Equation (17.1.9) coincides with the general Structure Equation (16.3.19). Since θ is a horizontal form, (17.1.10) follows immediately from formula (16.7.7) with $\sigma = \sigma_n^0$. ■

Remark 17.1.7. Using $\omega \wedge \theta(X, Y) = \omega(X)\theta(Y) - \omega(Y)\theta(X)$, the Structure Equations can also be written as

$$d\omega = -\omega \wedge \omega + \Omega, \quad d\theta = -\omega \wedge \theta + \Theta. \quad (17.1.11)$$

If we decompose the above forms with respect to the standard bases $\{\mathbf{e}_i\}$ in \mathbb{R}^n and $\{E^i_j\}$ in $\mathfrak{gl}_n(\mathbb{R})$,

$$\theta = \theta^i \mathbf{e}_i, \quad \Theta = \Theta^i \mathbf{e}_i, \quad \omega = \omega^i_j E^j_i, \quad \Omega = \Omega^i_j E^j_i, \quad (17.1.12)$$

then we obtain the Structure Equations in the form

$$d\omega^i_j = -\omega^i_k \wedge \omega^k_j + \Omega^i_j, \quad d\theta^i = -\omega^i_j \wedge \theta^j + \Theta^i. \quad (17.1.13)$$

The Bianchi identity for the curvature has a counterpart for the torsion.

Proposition 17.1.8 (Bianchi Identities [17, Prop. 2.1.13]). *Let ω, Ω and Θ be, respectively, the connection, curvature and torsion forms of a linear connection \mathcal{HM} . Then,*

$$D^\omega \Omega = 0, \quad (17.1.14)$$

$$D^\omega \Theta = \Omega \wedge \theta. \quad (17.1.15)$$

Proof. Equation (17.1.14) coincides with the general Bianchi identity (16.7.23). Equation (17.1.15) is an immediate consequence of Proposition 16.7.9 with $\sigma = \sigma_n^0$. It can also be verified directly by differentiating the first Structure Equation in the form (17.1.13). ■

Remark 17.1.9. The 1-forms ω and θ may be combined to make the so called *affine connection*, defined as the direct sum

$$\eta = \omega + \theta \in \Omega^1(\text{Fr } M, \mathfrak{gl}_n(\mathbb{R}) \oplus \mathbb{R}^n). \quad (17.1.16)$$

Clearly, $\mathfrak{a}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R}) \oplus \mathbb{R}^n$ is the Lie algebra of the affine group $\text{Aff}_n(\mathbb{R})$, with the additional commutation relations

$$[A, \mathbf{x}] = A\mathbf{x}, \quad [\mathbf{x}, \mathbf{y}] = 0, \quad A \in \mathfrak{gl}_n(\mathbb{R}), \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (17.1.17)$$

The form η is equivariant with respect to the direct sum representation $\text{Ad} \oplus \sigma_n^0$ of $\text{GL}_n(\mathbb{R})$. Accordingly, we may pass to the principal bundle $\text{Fr}_{\text{Aff}} M$ of *affine frames*. Then η is a principal connection on $\text{Fr}_{\text{Aff}} M$. Obviously,

$$D_\eta \eta = d\eta + \frac{1}{2}[\eta, \eta] = \Omega + \Theta, \quad (17.1.18)$$

that is, curvature and torsion constitute a joint object on $\text{Fr}_{\text{Aff}} M$, namely the curvature of η . Since $\ker \theta = \mathcal{V} \text{Fr } M$ and $\omega|_{\mathcal{V} \text{Fr } M}$ is injective, we see that for each $u \in \text{Fr } M$, the restriction of η to $T_u \text{Fr } M$ is injective and hence a linear isomorphism. In other words, η defines a global trivialization of $T \text{Fr } M$, which we will exhibit explicitly in Remark 17.2.7.

Since Ω and Θ are horizontal 2-forms on $\text{Fr } M$ of type Ad , they uniquely correspond to 2-forms on M with values in $\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \text{End}(\mathbb{R}^n)$ and $\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n$, respectively. Via the soldering isomorphism ϑ these can be turned into regular 2-forms on M . Recall also that $\theta_{\pi(u)} = \iota_u \circ \vartheta_u$, so $(d^\omega \vartheta)_{\pi(u)} = \iota_u \circ (D^\omega \theta)_u$. We thus have the definitions

$$\mathsf{T}_m(X_1, X_2) := \iota_u^\vartheta \circ \Theta_u(Y_1, Y_2) = \vartheta^{-1} \circ (d^\omega \vartheta)_m(X_1, X_2). \quad (17.1.19)$$

$$\mathsf{R}_m(X_1, X_2) := \iota_u^\vartheta \circ \Omega_u(Y_1, Y_2) \circ (\iota_u^\vartheta)^{-1} = \vartheta^{-1} \circ \mathsf{R}_m^\nabla(X_1, X_2) \circ \vartheta, \quad (17.1.20)$$

where $\pi(Y_i) = X_i$, $i = 1, 2$, and the 2-form $\mathsf{R}^\nabla \in \Omega^2(M; \text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n)$ was defined in (16.8.28). Since T takes values in TM and R in $TM \otimes T^*M$, they are actually regular tensor fields on M .

Definition 17.1.10 (Curvature and torsion tensor fields). The 2-forms $\mathsf{R} \in \Gamma_3^1(M)$ and $\mathsf{T} \in \Gamma_2^1(M)$ on M defined above are called the curvature and the torsion tensor fields of the linear connection ω , respectively.

The principal connection ω on $\text{Fr } M$ induces a covariant derivative operator ∇^ω on sections of $\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n$. Via the soldering isomorphism ϑ , this turns into a covariant derivative $\nabla^\omega : \Gamma^\infty(TM) \rightarrow \Gamma^\infty(T^*M \otimes TM)$:

$$(\nabla^\omega Y)_m(X) := \iota_u^\vartheta \circ (\widetilde{D^\omega \vartheta \circ Y})_u(X^*), \quad \pi(u) = m, \quad (17.1.21)$$

where $\widetilde{\vartheta \circ Y} \in \text{Hom}_{\text{GL}_n(\mathbb{R})}(\text{Fr } M, \mathbb{R}^n)$ is the usual equivariant function defined equal to $(\iota_u^\vartheta)^{-1} \circ Y(m)$. For each $X \in \mathfrak{X}(M)$ we get the associated operator

$$\nabla_X^\omega : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M), \quad \nabla_X^\omega Y := (\nabla^\omega Y)(X). \quad (17.1.22)$$

Once again, we shall assume that a connection has been chosen and write ∇ instead of ∇^ω . In fact, the connection can be recovered from the operator ∇ (since ∇ determines which vector fields, and hence frames, are parallel), which is why ∇ itself is often referred to as the “connection”.

Remark 17.1.11. 1. By (16.8.5), the formula for ∇ can be rewritten as $(\nabla_X Y)(m) = \iota_u^\vartheta(X_u^h(\widetilde{\vartheta \circ Y}))$, where X^h is the horizontal lift of X to $\text{Fr } M$. Thus, using $\theta_u(Y^*) = (\iota_u^\vartheta)^{-1} \circ \pi_*(Y^*) = (\iota_u^\vartheta)^{-1}(Y_m) = (\widetilde{\vartheta \circ Y})_u$, we obtain

$$(\nabla_X Y)_m = \iota_u^\vartheta(X_u^h(\theta(Y^*))). \quad (17.1.23)$$

2. Clearly ∇_X has all the properties listed in Proposition 16.8.10. Moreover, it induces covariant derivatives on all tensor bundles $T_l^k M$. A general formula is easily derived from (16.7.30) by taking σ for the tensor product representation of k copies of σ_n^0 and l copies of its dual.

The following is a simple exercise, similar to Willmore’s Theorem 9.3.10.

Proposition 17.1.12 ([17, Prop. 2.1.18]). *Let \mathcal{HM} be a linear connection on a manifold M and let ∇ be its covariant derivative on TM . Then the operator $\nabla_X : \Gamma_s^r(M) \rightarrow \Gamma_s^r(M)$ on tensor fields of type (r, s) is uniquely determined by the following properties.*

1. $\nabla_X f = X(f)$ for $f \in C^\infty(M)$.
2. ∇_X is a derivation of the tensor algebra.
3. ∇_X commutes with tensor contractions.

The curvature tensor is expressed in terms of covariant derivatives as before, but now we can obtain a similar expression for the torsion tensor.

Proposition 17.1.13 ([17, Prop. 2.1.19]). *Let ∇ be the covariant derivative of a linear connection \mathcal{HM} on M . Then its curvature and torsion tensor fields are given by*

$$\mathbf{R}(X, Y) = [\nabla_X, \nabla_Y] - \nabla_{[X, Y]}, \quad (17.1.24)$$

$$\mathbf{T}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]. \quad (17.1.25)$$

Proof. The first formula follows from Proposition 16.8.13 as a special case. To prove the second formula, we use the definition of the covariant exterior derivative to write $\Theta(X^h, Y^h) = d\theta(X^h, Y^h)$. Using this, together with (17.1.23), and $\pi_*([X^h, Y^h]) = [X, Y]$,

we obtain

$$\mathsf{T}(X, Y)_m = \iota_u^\theta \circ \Theta_u(X^h, Y^h) = \quad (17.1.26)$$

$$= \iota_u^\theta \left(X_u^h(\theta(Y^h)) - Y_u^h(\theta(X^h)) - \theta_u([X^h, Y^h]) \right) = \quad (17.1.27)$$

$$= (\nabla_X Y - \nabla_Y X - [X, Y])_m. \quad (17.1.28)$$

■

Finally, ∇ induces parallel transports on all tensor bundles $T_s^r M$, and we obtain corresponding holonomy groups. The Ambrose-Singer Theorem 16.5.8 then provides a deep relation between holonomy and curvature, which has tremendous consequences for the structure theory of (pseudo-)Riemannian manifolds.

Remark 17.1.14. Note that if $H = \mathrm{GL}_n(\mathbb{R})$ is the subgroup of $G = \mathrm{Aff}_n(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R}) \rtimes \mathbb{R}^n$ corresponding to the first factor, then since the second factor \mathbb{R}^n is a normal subgroup, $\mathrm{Ad}(H)$ leaves it invariant. Therefore the decomposition of the Lie algebra $\mathfrak{a}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R}) \oplus \mathbb{R}^n$ is $\mathrm{Ad}(H)$ -invariant. In other words, (G, H) is a reductive Klein geometry on the affine space $\mathrm{A}(\mathbb{R}^n)$.

The natural projection $\pi : G \rightarrow G/H \cong \mathbb{R}^n$ is a principal $\mathrm{GL}_n(\mathbb{R})$ -bundle that carries the Maurer-Cartan form $\Theta_G \in \Omega^1(\mathrm{Aff}_n(\mathbb{R}), \mathfrak{a}_n(\mathbb{R}))$. Now we may split $\Theta_G = \omega + \theta$ according to the above reductive decomposition of $\mathfrak{a}_n(\mathbb{R})$. Since the splitting is $\mathrm{GL}_n(\mathbb{R})$ -invariant, both ω and θ are $\mathrm{GL}_n(\mathbb{R})$ -equivariant forms. The form θ associates to each element $g \in \mathrm{Aff}_n(\mathbb{R})$ a linear isomorphism $T_{\pi(g)}\mathrm{A}(\mathbb{R}^n) \rightarrow \mathbb{R}^n$, and thus identifies $\mathrm{Aff}_n(\mathbb{R})$ with the frame bundle $\mathrm{Fr}\mathrm{A}(\mathbb{R}^n)$. Hence, the component ω may be viewed as a linear connection on $\mathrm{A}(\mathbb{R}^n)$ and one immediately sees that this gives the canonical flat connection on $\mathrm{A}(\mathbb{R}^n)$.

Remark 17.1.15. Returning to arbitrary manifolds M , we may now view the data defining a linear connection on M as an analog of the Klein geometry $\mathrm{A}(\mathbb{R}^n)$. Indeed, the analog of the principal $\mathrm{GL}_n(\mathbb{R})$ -bundle $\mathrm{Aff}_n(\mathbb{R}) \rightarrow \mathrm{A}(\mathbb{R}^n)$ is the frame bundle $\mathrm{Fr}M \rightarrow M$. On the other hand, the affine connection form $\eta = \omega + \theta$ may be viewed as an element of $\Omega^1(\mathrm{Fr}M, \mathfrak{a}_n(\mathbb{R}))$, which is an analog of the Maurer-Cartan form on $\mathrm{A}(\mathbb{R}^n)$. In fact, we have already noted above that η is $\mathrm{GL}_n(\mathbb{R})$ -equivariant, reproduces generators of Killing vector fields, and trivializes TM . More formally, we have:

1. $\Phi_g^* \eta = \mathrm{Ad}_g^{-1} \circ \eta$, $g \in \mathrm{GL}_n(\mathbb{R})$,
2. $\eta(A_*) = A$ for all $X \in \mathfrak{gl}_n(\mathbb{R}) \subset \mathfrak{a}_n(\mathbb{R})$,
3. $\eta_u : T_u \mathrm{Fr}M \rightarrow \mathfrak{a}_n(\mathbb{R})$ is a linear isomorphism for any $u \in \mathrm{Fr}M$.

These are exactly the strongest analogs of the properties of the Maurer-Cartan form that make sense on a general principal $\mathrm{GL}_n(\mathbb{R})$ -bundle.

To complete this interpretation, observe the if we have an arbitrary principal $\mathrm{GL}_n(\mathbb{R})$ -bundle

$P \xrightarrow{\pi} M$ with a one-form $\eta \in \Omega^1(P, \mathfrak{a}_n(\mathbb{R}))$ which satisfies the three properties above. It can be uniquely decomposed as a sum $\eta = \omega + \theta$ where ω is $\mathrm{gl}_n(\mathbb{R})$ -valued and θ is \mathbb{R}^n -valued. Then properties 2 and 3 ensure that $\ker \theta$ is exactly $\mathcal{V}P$. Hence θ_u descends to a linear isomorphism $T_u P / \mathcal{V}_u P \cong T_{\pi(u)} M \rightarrow \mathbb{R}^n$, i.e. a frame on $T_{\pi(u)} M$. This defines a morphism of principal bundles $P \rightarrow \mathrm{Fr} M$, and property 1 then ensures that this is an isomorphism. By construction, the pullback of the soldering form on $\mathrm{Fr} M$ along this isomorphism is exactly θ . And as before, ω defines a principal connection on $\mathrm{Fr} M$.

This shows that a linear connection on M can be equivalently described as an *affine structure*, i.e. a principal $\mathrm{GL}_n(\mathbb{R})$ -bundle $P \rightarrow M$ together with a 1-form η which has properties 1-3 from above. The basic features of linear connections on manifolds can be nicely phrased in this picture, which motivates many developments in general Cartan geometries.

Example 17.1.16 (Canonical connections on a Lie group, part 4). We continue Example 16.8.15 and compute the torsion form of ω^λ . First, since we are not dealing with a regular frame bundle, we need to pick a soldering form on P . This needs to be a \mathfrak{g} -valued horizontal 1-form of type Ad whose induced bundle morphism $TM \rightarrow P \times_G \mathfrak{g}$ is an isomorphism. Luckily we have already constructed exactly such a form θ in Example 16.2.6, given by $\theta = \mathrm{pr}_1^* \Theta_G - \mathrm{pr}_2^* \Theta_G$. Using the formula for the covariant exterior derivative from Proposition 16.7.3 in conjunction with the expressions we have obtained for ω^λ and θ , and also using the fact that the wedge product of \mathfrak{g} -valued 1-forms is commutative, we find

$$\Theta = D^\lambda \theta = d\theta + \omega^\lambda \wedge \theta = \lambda[\theta, \theta]. \quad (17.1.29)$$

This also agrees with the expression for the torsion tensor on invariant vector fields, which can be obtained from (17.1.25):

$$\mathsf{T}^\lambda(A_L, B_L) = \lambda[A, B]_L, \quad \mathsf{T}^\lambda(A_R, B_R) = \lambda[A, B]_R. \quad (17.1.30)$$

Note that the sign doesn't change due to the identity $[A, B]_R = -[A_R, B_R]$, but this is expected since T is a tensor of type $(1, 2)$. As a TG -valued 2-form, this tensor can be written as

$$\mathsf{T}_g^\lambda = \frac{\lambda}{2} L_{g*} [\Theta_G, \Theta_G]. \quad (17.1.31)$$

Due to these values of the torsion, ∇^{+1} , ∇^0 , and ∇^{-1} are known as (+)-, 0-, and (-)-connections, respectively. Moreover, by the Bianchi identity for the torsion, $\nabla^{\pm 1} T^{\pm 1} = 0$, so the torsion tensor itself is parallel for $\lambda = \pm 1$, i.e. its components in the respective global frame are constant and equal to the structure constants of G , up to a sign. As we will see, these data are enough to locally characterize the Lie group structure as a special kind of G -structure.

Explaining the meaning of torsion is especially easy on Lie groups due to the presence of canonical flat (\pm)-connections whose torsion is $\mathsf{T}(A_L, B_L) = \pm[A, B]_L$. Suppose $\gamma : I \rightarrow G$

is a smooth path. Then $t \mapsto \dot{\gamma}(t)$ is a smooth path in TG . We can use left translations to turn this into a path in $T_e G = \mathfrak{g}$:

$$\delta\gamma(t) = L_{\gamma(t)*}^{-1}\dot{\gamma}(t) \in \mathfrak{g}. \quad (17.1.32)$$

Note that left translations coincide with parallel transport under the $(-)$ -connection, so $L_{\gamma(t)*}^{-1} = (\text{tra}_\gamma^{(-)}(t))^{-1}$. Now we can integrate $\delta\gamma(t)$ to form another path $\tilde{\gamma} : I \rightarrow T_e G$:

$$\tilde{\gamma}^-(t) = \iint_0^t \delta\gamma(s) ds \in \mathfrak{g}. \quad (17.1.33)$$

The natural question now is what $\tilde{\gamma}^-(1)$ is. Suppose γ is a loop. Since we can also write $\delta\gamma(t) = \Theta_G(\dot{\gamma}(t))$, we have

$$\tilde{\gamma}^-(1) = \int_0^1 \Theta_G(\dot{\gamma}(s)) ds = \oint_\gamma \Theta_G. \quad (17.1.34)$$

Now suppose γ is nullhomotopic and S is a 2-dimensional submanifold of G such that $\partial S = \text{im}(\gamma)$. Then, using the Stokes' theorem and the Maurer-Cartan equation, we have

$$\tilde{\gamma}^-(1) = -\frac{1}{2} \int_S [\Theta_G, \Theta_G] = \int_S \Theta_G \circ \mathsf{T}^-. \quad (17.1.35)$$

This formula is also true for T^+ if $\tilde{\gamma}^+$ is defined using right translations. Here, θ_G merely acts so that the values of the integrand belong to one vector space (\mathfrak{g}). In local coordinates, this integral is proportional to the value of T^- at the identity times the area of S . This result, in fact, holds for arbitrary affine connections (although in general there is no analog of Θ_G that can transport all values of T to one point, so the general formula can only be written in local coordinates). Therefore, torsion can be interpreted as the “displacement of the origin” of a tangent space as it is being *rolled without slipping* along an infinitesimal loop (this is easier to imagine as M rolling on \mathbb{R}^n so that the point of contact follows γ). For this reason the soldering form θ is also sometimes called the *displacement form*.

17.2 Geodesics

What distinguishes affine parallel transport from parallel transport on general bundles is that everything happens on the base manifold M , so there is a special class of distinguished curves that we discuss now. As before, for any smooth path $\gamma : I_\epsilon \rightarrow M$ we have a unique parallel transport of tangent vectors along γ . Let $\dot{\gamma}$ be the tangent vector field of γ :

$$\dot{\gamma}(t) = \gamma_{*t} \left(\partial_t \Big|_t \right). \quad (17.2.1)$$

Applying the notions developed in the previous section, a vector field $X \in \mathfrak{X}(M)$ is parallel along γ iff

$$\nabla_{\dot{\gamma}} X := \nabla_{\partial_t}^{\gamma} X = 0, \quad (17.2.2)$$

where ∇^{γ} is the covariant derivative along the map γ and X must be viewed as a section of TM along γ (that is, we should really write $X \circ \gamma$ instead of X). In particular, since $\dot{\gamma}$ is a section of TM along γ , we may consider the equation

$$\nabla_{\dot{\gamma}} \dot{\gamma}(t) = \nabla_{\partial_t}^{\gamma} \dot{\gamma} = 0. \quad (17.2.3)$$

Definition 17.2.1 (Geodesic). Let \mathcal{HM} be a linear connection. A path $\gamma : I_{\epsilon} \rightarrow M, t \mapsto \gamma(t)$ is called a geodesic with respect to \mathcal{HM} if it fulfills the *geodesic equation* $\nabla_{\dot{\gamma}} \dot{\gamma}(t) = 0$.

The following proposition states that the set of geodesics is invariant under affine reparametrizations. This follows immediately from the fact that rescaling the variable t is equivalent to rescaling $\dot{\gamma}$.

Proposition 17.2.2 ([17, Prop. 2.1.21]). *If $\gamma : I_{\epsilon} \rightarrow M$ is a geodesic, then for any $\alpha, \beta \in \mathbb{R}$ the curve $t \mapsto \gamma(\alpha t + \beta)$ is a geodesic, too.*

The following vector fields on the frame bundle will turn out to be closely connected to geodesics. Given a linear connection on M , corresponding to a principal connection ω on $\text{Fr } M$, any $\mathbf{x} \in \mathbb{R}^n$ defines a horizontal vector field $B(\mathbf{x})$ on $\text{Fr } M$ by assigning to $u \in \text{Fr } M$ the unique horizontal lift of $\iota_u^{\theta}(\mathbf{x}) \in T_{\pi(u)} M$ to the point u .

Definition 17.2.3 (Standard vector fields). The vector field $B(\mathbf{x})_u = (\iota_u^{\theta}(\mathbf{x}))_u^h$ is called the horizontal standard vector field defined by $\mathbf{x} \in \mathbb{R}^n$.

Proposition 17.2.4 ([17, Prop. 2.1.7]). *For any $\mathbf{x} \in \mathbb{R}^n$, the horizontal standard vector field satisfies:*

1. $\theta(B(\mathbf{x})) = \mathbf{x}$,
2. $\Phi_{g*} B(\mathbf{x}) = B(g^{-1}\mathbf{x})$, $g \in \text{GL}_m(\mathbb{R})$,
3. if $\mathbf{x} \neq 0$, then $B(\mathbf{x})$ vanishes nowhere.

Proof. 1. We calculate

$$\theta_u(B(\mathbf{x})) = (\iota_u^{\theta})^{-1}(\pi_*(B(\mathbf{x})_u)) = (\iota_u^{\theta})^{-1}(\iota_u^{\theta}(\mathbf{x})) = \mathbf{x}. \quad (17.2.4)$$

2. By Proposition 17.1.2 and point 1,

$$\theta(\Phi_{g*} B(\mathbf{x})) = \Phi_g^* \theta(B(\mathbf{x})) = g^{-1} \theta(B(\mathbf{x})) = g^{-1} \mathbf{x}, \quad (17.2.5)$$

and thus $\pi_*(\Phi_{g*}B(\mathbf{x})) = \iota_u^\theta(g^{-1}\mathbf{x})$. Since $\Phi_{g*}B(\mathbf{x})$ is horizontal, the assertion follows from the uniqueness of horizontal lifts.

3. Clearly, $B(\mathbf{x})_u = 0$ iff $\iota_u^\theta(\mathbf{x}) = 0$, and thus iff $\mathbf{x} = 0$, since ι_u^θ is an isomorphism. ■

Remark 17.2.5. Let $\{\mathbf{e}_i\}$ be the standard basis in \mathbb{R}^n . Then, the horizontal standard vector fields $B_i = B(\mathbf{e}_i)$ span the horizontal distribution defined by $\mathcal{H}M$. Moreover, $B(\mathbf{x})$ is uniquely determined by the conditions

$$\theta(B(\mathbf{x})) = \mathbf{x}, \quad \omega(B(\mathbf{x})) = 0. \quad (17.2.6)$$

Lemma 17.2.6 ([17, Lem. 2.1.9]). *Let A_* be the Killing vector field on $\text{Fr } M$ generated by $A \in \mathfrak{gl}_n(\mathbb{R})$ and let $\mathbf{x} \in \mathbb{R}^n$. Then*

$$[A_*, B(\mathbf{x})] = B(A\mathbf{x}). \quad (17.2.7)$$

Proof. Let $g_t = e^{tA}$. Using Proposition 17.2.4(2), we obtain

$$[A_*, B(\mathbf{x})] = (\mathcal{L}_{A_*} B(\mathbf{x}))_u = \frac{d}{dt} \Big|_0 \left(\Phi_{g_t*}^{-1} B(\mathbf{x}) \right) = \frac{d}{dt} \Big|_0 B(g_t \mathbf{x})_u = B(A\mathbf{x})_u. \quad (17.2.8)$$

Remark 17.2.7. In Remark 17.1.9 we observed that the affine connection form $\eta = \omega + \theta$ trivializes the tangent bundle of $\text{Fr } M$. Let us now present a natural global frame. Let $\{\mathbf{e}_i\}$ and $\{E^j_i\}$ be the standard bases of \mathbb{R}^n and $\mathfrak{gl}_n(\mathbb{R})$, respectively. Let B_i be the horizontal standard vector field with respect to a chosen connection $\mathcal{H}E$ generated by \mathbf{e}_i and let E^j_{i*} be the Killing vector field generated by E^j_i . Since $\{E^j_{i*}\}$ span the vertical subspace $\mathcal{V}_u \text{Fr } M$ at every $u \in \text{Fr } M$, and since $\{B_i\}$ span the horizontal subspace $\mathcal{H}_u \text{Fr } M$, these $n^2 + n$ vector fields provide a global frame for $T \text{Fr } M$, which is, therefore, trivial. Thus the manifold $\text{Fr } M$ carries a *absolute parallelism* given by these vector fields. Moreover, the vector fields B_i, E^j_{i*} are dual to the 1-forms θ^i, ω^i_j :

$$\theta^k(B_i) = \delta_i^k, \quad \theta^k(E^j_{i*}) = 0, \quad (17.2.9)$$

$$\omega^k{}_l(B_i) = 0, \quad \omega^k{}_l(E^j_{i*}) = \delta^k{}_i \delta^j{}_l. \quad (17.2.10)$$

Thus, the absolute parallelism on $T^* \text{Fr } M$ is provided by the 1-forms θ^i, ω^i_j . In more abstract terms, the affine connection form η induces an absolute parallelism on $\text{Fr}_{\text{aff}} M$. As a consequence, every horizontal k -form α on $\text{Fr } M$ may be expanded with respect to the 1-forms θ^i :

$$\alpha = \frac{1}{k!} \alpha_{i_1 \dots i_k} \theta^{i_1} \wedge \dots \wedge \theta^{i_k}. \quad (17.2.11)$$

In particular,

$$\Omega^i{}_j = \frac{1}{2} \Omega_{klj}^i \theta^k \wedge \theta^l, \quad \Theta^i = \frac{1}{2} \Theta_{jk}^i \theta^j \wedge \theta^k. \quad (17.2.12)$$

Remark 17.2.8. Since, for any $u \in \text{Fr } M$, the assignment $\mathbb{R}^n \rightarrow \mathcal{H}_u(\text{Fr } M)$, $\mathbf{x} \mapsto B(\mathbf{x})$, is an isomorphism of vector spaces, we have an induced isomorphism

$$b(u) : \bigwedge^2 \mathbb{R}^n \rightarrow \bigwedge^2 \mathcal{H}_u(\text{Fr } M), \quad b(u)(\mathbf{x} \wedge \mathbf{y}) = B(\mathbf{x})_u \wedge B(\mathbf{y})_u. \quad (17.2.13)$$

Using this, we get yet another useful presentation of curvature and torsion. We define maps

$$\mathcal{R} : \text{Fr } M \rightarrow \bigwedge^2 (\mathbb{R}^n)^* \otimes \mathfrak{gl}_n(\mathbb{R}), \quad \mathcal{T} : \text{Fr } M \rightarrow \bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n \quad (17.2.14)$$

by

$$\mathcal{R}(u) := \Omega_u \circ b(u), \quad \mathcal{T}(u) := \Theta_u \circ b(u). \quad (17.2.15)$$

In the sequel, \mathcal{R} and \mathcal{T} will be referred to as the *curvature and torsion maps*. Note also that $b(u)^{-1}(X_u \wedge Y_u) = \theta_u(X_u) \wedge \theta_u(Y_u)$. Using horizontality of Θ and Ω , together with (14.2.13), we find:

$$\mathcal{R}(\Phi_g(u))(\mathbf{x}, \mathbf{y}) = \text{Ad}_g^{-1} \circ (\mathcal{R}(u)(g\mathbf{x}, g\mathbf{y})), \quad (17.2.16)$$

$$\mathcal{T}(\Phi_g(u))(\mathbf{x}, \mathbf{y}) = g^{-1} \circ (\mathcal{T}(u)(g\mathbf{x}, g\mathbf{y})). \quad (17.2.17)$$

By Proposition 14.2.17, to \mathcal{R} and \mathcal{T} , there correspond unique sections of the associated bundles

$$\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \left(\bigwedge^2 (\mathbb{R}^n)^* \otimes \mathfrak{gl}_n(\mathbb{R}) \right), \quad \text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \left(\bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n \right), \quad (17.2.18)$$

respectively. By (17.2.15), they are given by

$$m \mapsto \iota_u^\vartheta \circ \mathcal{R}(u) \circ (\iota_u^\vartheta)^{-1} = R_m \circ \bigwedge^2 \iota_u^\vartheta, \quad m \mapsto \iota_u^\vartheta \circ \mathcal{T}(u) = T_m \circ \bigwedge^2 \iota_u^\vartheta, \quad (17.2.19)$$

where $\bigwedge^2 \iota_u^\vartheta : \mathbb{R}^n \wedge \mathbb{R}^n \rightarrow T_m M \wedge T_m M$ and $m = \pi(u)$.

Proposition 17.2.9 ([17, Prop. 2.1.22]). *Let $\mathcal{H}M$ be a linear connection on M . Then the projection, under $\pi : \text{Fr } M \rightarrow M$, of any integral curve of a standard vector field is a geodesic. Conversely, every geodesic is obtained this way.*

Proof. Let $\mathbf{x} \in \mathbb{R}^n$. By definition, $B(\mathbf{x})_u$ is the unique $\mathcal{H}M$ -horizontal lift of $\iota_u^\vartheta(\mathbf{x}) \in T_{\pi(u)}M$ to $u \in \text{Fr } M$. Let $t \mapsto \tilde{\gamma}(t)$ be an integral curve of $B(\mathbf{x})$. Define $\gamma := \pi \circ \tilde{\gamma}$. Then, using the natural isomorphism $\vartheta : TM \rightarrow \text{Fr } M \times_{\text{GL}_n(\mathbb{R})} \mathbb{R}^n$, and defining $\iota_x^\vartheta := \vartheta^{-1} \circ \iota_x : \text{Fr } M \rightarrow TM$,

$$\dot{\gamma}(t) = \pi_* \circ \dot{\tilde{\gamma}}(t) = \pi_*(B(\mathbf{x})_{\tilde{\gamma}(t)}) = \iota_{\tilde{\gamma}(t)}^\vartheta(\mathbf{x}) = \iota_{\tilde{\gamma}(t)}^\vartheta(\tilde{\gamma}(t)). \quad (17.2.20)$$

Thus, by (16.9.19) and (16.4.3), we have

$$\nabla_{\dot{\gamma}} \dot{\gamma}(t) = \omega^E \left((\iota_{\tilde{\gamma}(t)}^\vartheta)_* (\dot{\tilde{\gamma}}(t)) \right) = 0. \quad (17.2.21)$$

Conversely, let γ be a geodesic and let $u_0 \in \text{Fr } M$ be such that $\pi(u_0) = \gamma(0)$ and let $\mathbf{x} := (\iota_{u_0}^\vartheta)^{-1}(\dot{\gamma}(0)) \in \mathbb{R}^n$. Let $t \mapsto \tilde{\gamma}(t)$ be the horizontal lift of γ through u_0 . If $\mathbf{x} = 0$, we are done. Thus, let $\mathbf{x} \neq 0$. Then there exists a curve $t \mapsto \sigma(t)$ in $\text{Fr } M$ such that $\dot{\gamma}(t) = \iota_{\sigma(t)}^\vartheta(\mathbf{x})$. Hence

$$\frac{d}{dt} \dot{\gamma}(t) = (\iota_{\tilde{\gamma}(t)}^\vartheta)_* \dot{\sigma}(t). \quad (17.2.22)$$

Since γ is a geodesic, that is, $\frac{d}{dt} \dot{\gamma}(t) \in T(TM)$, this formula implies that $t \mapsto \sigma(t)$ is horizontal in $\text{Fr } M$. Since $\sigma(0) = u_0$ and $\pi \circ \sigma = \gamma$, uniqueness of the horizontal lift implies $\sigma = \tilde{\gamma}$. Thus, $\dot{\gamma}(t) = \iota_{\tilde{\gamma}(t)}^\vartheta(\mathbf{x})$ and, since $\tilde{\gamma}$ is horizontal,

$$\theta(\dot{\gamma}(t)) = (\iota_{\tilde{\gamma}(t)}^\vartheta)^{-1} \circ \pi_* \circ \dot{\tilde{\gamma}}(t) = (\iota_{\tilde{\gamma}(t)}^\vartheta)^{-1}(\dot{\gamma}(t)) = \mathbf{x}. \quad (17.2.23)$$

Thus $t \mapsto \tilde{\gamma}(t)$ is an integral curve of $B(\mathbf{x})$. ■

Corollary 17.2.10 ([17, Cor. 2.1.23]). *Let $\mathcal{H}M$ be a linear connection on M . For every $m \in M$ and every $X \in T_m M$, there exists a unique geodesic $\gamma : I_\epsilon \rightarrow M$ with initial conditions (m, X) , that is, $\gamma(0) = m$ and $\dot{\gamma}(0) = X$.*

We say that a linear connection $\mathcal{H}M$ on M is complete if every geodesic may be extended to $I_\epsilon = \mathbb{R}$.

Corollary 17.2.11 ([17, Cor. 2.1.24]). *A linear connection on M is complete iff every horizontal standard vector field on $\text{Fr } M$ is complete.*

Definition 17.2.12 (Exponential map). Let $\mathcal{H}M$ be a complete linear connection on a manifold M . The exponential map $\exp : TM \rightarrow M$ is defined for every $m \in M, X \in T_m M$ by taking the unique geodesic γ_X with initial conditions $\gamma(0) = m, \dot{\gamma}(0) = X$ and putting

$$\exp(X) := \gamma_X(1). \quad (17.2.24)$$

Remark 17.2.13. The exponential map may be well-defined even if the connection is not complete. In this case, \exp is defined on a neighborhood of the zero section in TM . For every $m \in M$, it yields a local diffeomorphism from a 0-neighborhood in $T_m M$ onto a neighborhood U_m of m in M .

Proposition 17.2.14 ([47, p. 249]). *Let ∇ and $\tilde{\nabla}$ be two affine connections on M . TFAE:*

- (a) ∇ and $\tilde{\nabla}$ have the same geodesics,
- (b) $\tilde{\nabla}_X X = \nabla_X X$ for all $X \in \mathfrak{X}(M)$,
- (c) $\tilde{\nabla}_X Y + \tilde{\nabla}_Y X = \nabla_X Y + \nabla_Y X$ for all $X, Y \in \mathfrak{X}(M)$.

Proof. (a) \Rightarrow (b): let $X_m \in T_m M$ be non-zero and let γ_{X_m} be the corresponding geodesic. Let X be a vector field in a neighborhood of m that extends $\dot{\gamma}$ along a part of γ . Then

$$\tilde{\nabla}_X X(m) = \tilde{\nabla}_{X_m} X = \tilde{\nabla}_{\dot{\gamma}(0)} X = \nabla_{\dot{\gamma}(0)} X = \nabla_X X(m). \quad (17.2.25)$$

(b) \Rightarrow (a): Let $\gamma : I_\epsilon \rightarrow M$ be a geodesic for ∇ and let $X \in \mathfrak{X}(M)$ be an extension of $\dot{\gamma}$. Then

$$\tilde{\nabla}_{\dot{\gamma}} X = \nabla_{\dot{\gamma}} X, \quad (17.2.26)$$

thus γ is a geodesic for $\tilde{\nabla}$.

(b) \Leftrightarrow (c): Define $S(X, Y) = (\tilde{\nabla}_X Y + \tilde{\nabla}_Y X) - (\nabla_X Y + \nabla_Y X)$. Then $S(X, Y) = S(Y, X)$ and therefore $S(X + Y, X + Y) - S(X, X) - S(Y, Y) = 2S(X, Y)$, so the vanishing of S on the diagonal, (b), is equivalent to the vanishing of S everywhere, (c). ■

Corollary 17.2.15 ([47, p. 250]). *Two affine connections ∇ and $\tilde{\nabla}$ on M coincide iff both their geodesics and torsion tensors coincide.*

Proof. Denoting the difference of the antisymmetric parts of the connections

$$A(X, Y) = (\tilde{\nabla}_X Y - \tilde{\nabla}_Y X) - (\nabla_X Y - \nabla_Y X), \quad (17.2.27)$$

we have $A(X, Y) = \tilde{T}(X, Y) - T(X, Y)$. By the proposition above, the two connections with the same geodesics coincide iff $A = 0$, i.e. $\tilde{T} = T$. ■

Corollary 17.2.16 ([47, p. 250]). *For every affine connection $\tilde{\nabla}$ on M there is a unique torsion-free connection ∇ with the same geodesics.*

Proof. Uniqueness follows from the previous corollary. For existence, define $\nabla_X Y = \tilde{\nabla}_X Y - \frac{1}{2}\tilde{T}(X, Y)$, checking easily that ∇ is a connection. Since \tilde{T} is skew-symmetric, $\tilde{\nabla}$ and ∇ have the same symmetric parts, i.e. $S = 0$, and thus have the same geodesics. Also $T = \tilde{T} - A = 0$, so ∇ is torsion-free. ■

Remark 17.2.17. This statement should not be confused with the existence of Levi-Civita connections. So far we have not imposed any compatibility constraints on the connections. In the presence of such constraints torsion-free connections may not exist at all. In the context of metric connections we will derive the existence of a unique torsion-free metric connection from general theory of G -structures in the next section.

Example 17.2.18 (Canonical connections on a Lie group, part 5). Picking up Example 17.1.16, recall that the soldering form on $P = G \times G$ was defined as $\theta = \text{pr}_1^* \Theta_G - \text{pr}_2^* \Theta_G$. Thus, given $A \in \mathfrak{g}$, the frame $(g_1, g_2) \in P$ maps A to the tangent vector

$$\iota_{(g_1, g_2)}^\vartheta(A) = \vartheta^{-1}([(g_1, g_2), A]) = L_{g_1 g_2^{-1}*} \text{Ad}_{g_2} A = L_{g_1*} R_{g_2*}^{-1} A. \quad (17.2.28)$$

Then the horizontal lift of this vector under the connection ω^λ is

$$B^\lambda(A)_{(g_1, g_2)} = (L_{g_1*} R_{g_2*}^{-1} A)_{(g_1, g_2)}^h = \frac{1-\lambda}{2} L_{g_1*} A \oplus \frac{-1-\lambda}{2} L_{g_2*} A \quad (17.2.29)$$

as an element of $T_{g_1}G \times T_{g_2}G$. Indeed, it is trivial to see that $\omega^\lambda(B^\lambda(A)) = 0$ and $\theta(B^\lambda(A)) = A$. In particular, $B^-(A)$ and $B^+(A)$ generate the flows $(g_1 e^{tA}, g_2)$ and $(g_1, e^{-tA} g_2)$, respectively. In either case, therefore, all geodesics on G have the form $g_1 e^{tA} g_2^{-1}$.

This allows us to solve for geodesics. It is easy to see that the curves

$$t \mapsto \left(g_1 \left(\exp \frac{1-\lambda}{2} t A \right), \exp \left(\frac{-1-\lambda}{2} t A \right) g_2 \right) \quad (17.2.30)$$

are integral curves of $B^\lambda(A)$ for all $(g_1, g_2) \in P$. In particular, all geodesics starting at the identity have the form $g e^{tA} g^{-1} = e^{t\text{Ad}_g A}$, i.e. all of them are OPSs. By Proposition 17.2.14, all of the connections ∇^λ have the same symmetric part and differ only in their torsion. We will examine the only torsion-free connection among them, ∇^0 , separately in Example 17.4.20.

17.3 Local formulas

Now we describe these structures locally. Let

$$m \mapsto \mathbf{e}(m) = (e_1(m), \dots, e_n(m)) \quad (17.3.1)$$

be a local section of $\text{Fr } M$, that is, a local frame of TM , and let

$$m \mapsto \alpha(m) = (\alpha^1(m), \dots, \alpha^n(m)) \quad (17.3.2)$$

be its dual coframe. Recall that $\iota_{\mathbf{e}(m)}^\vartheta(\mathbf{e}_i) = e_i(m)$, where $\{\mathbf{e}_i\}$ is the standard basis of \mathbb{R}^n .

Lemma 17.3.1 ([17, Lem. 2.1.26]). *For any local frame \mathbf{e} ,*

$$\mathbf{e}^*\theta = \alpha^i \otimes \mathbf{e}_i. \quad (17.3.3)$$

Proof. For any $X \in T_m M$, we calculate

$$(\mathbf{e}^*\theta)_m(X) = \theta_{\mathbf{e}(m)}(\mathbf{e}_*(X)) = (\iota_{\mathbf{e}(m)}^\vartheta)^{-1}(\pi_* \circ \mathbf{e}_*(X)) = (\iota_{\mathbf{e}(m)}^\vartheta)^{-1}(X). \quad (17.3.4)$$

Thus, decomposing $X = X^i e_i(m)$ and using $\iota_{\mathbf{e}(m)}^\vartheta(\mathbf{e}_i) = e_i(m)$, we obtain

$$(\mathbf{e}^*\theta)_m(X) = X^i(m) \mathbf{e}_i = \alpha_m^i(X) \mathbf{e}_i. \quad (17.3.5)$$

Thus, for components of θ with respect to the decomposition (17.1.12),

$$\mathbf{e}^*\theta^i = \alpha^i. \quad (17.3.6)$$

Next, the local representative $\mathcal{A} = \mathbf{e}^*\omega$ of the principal connection form ω with values in $\mathfrak{gl}_n(\mathbb{R})$ may be written as

$$\mathcal{A} = \mathcal{A}^i{}_k E^k{}_l = \Gamma^i{}_{jk} \alpha^j \otimes E^k{}_i. \quad (17.3.7)$$

The coefficient functions $\Gamma^i{}_{jk}$ are called the *Christoffel symbols* of \mathcal{HM} in the local frame \mathbf{e} . ■

Remark 17.3.2. Consider a change $\mathbf{e} \rightarrow \mathbf{e}'$ of the local frame (a passive gauge transformation). Using (16.3.34), we obtain the following induced transformation formula for the Christoffel symbols

$$\Gamma'^l{}_{mn} = \Gamma^i{}_{jk} \tau^j{}_m \tau^k{}_n (\tau^{-1})^l{}_i + \tau^j{}_m (\partial_j \tau^i{}_n) (\tau^{-1})^l{}_i. \quad (17.3.8)$$

Let us calculate the local representative of curvature and torsion. For that purpose, take the pullback of (17.2.12) along \mathbf{e} ,

$$\mathbf{e}^*\Omega^i{}_j = \frac{1}{2} (\mathbf{e}^*\Omega^i{}_{klj}) \alpha^k \wedge \alpha^l, \quad \mathbf{e}^*\Theta^i = \frac{1}{2} (\mathbf{e}^*\Theta^i{}_{jk}) \alpha^j \wedge \alpha^l, \quad (17.3.9)$$

and denote the local coefficient functions as follows

$$R^i{}_{klj} = \mathbf{e}^*\Omega^i{}_{klj}, \quad T^i{}_{jk} = \mathbf{e}^*\Theta^i{}_{jk}. \quad (17.3.10)$$

To calculate them we use the Structure Equations in the form (17.1.13). Taking the pullback of the first equation yields

$$\frac{1}{2} R^i{}_{klj} \alpha^k \wedge \alpha^l = d\mathcal{A}^i{}_j + \mathcal{A}^i{}_k \wedge \mathcal{A}^k{}_j. \quad (17.3.11)$$

Inserting (17.3.7) into this equation, we obtain

$$R^i_{\ jkl} = e_j(\Gamma^i_{\ kl}) - e_k(\Gamma^i_{\ jl}) + \Gamma^m_{\ kl}\Gamma^i_{\ jm} - \Gamma^m_{\ jl}\Gamma^i_{\ km} - C^m_{\ jk}\Gamma^i_{\ ml}, \quad (17.3.12)$$

where $C^i_{\ jk}$ are the structure functions of the local frame ϵ defined by

$$[e_j, e_k] = C^i_{\ jk}e_i. \quad (17.3.13)$$

In the same way, taking the pullback of the second equation in (17.1.13), we read off

$$T^i_{\ jk} = \Gamma^i_{\ jk} - \Gamma^i_{\ kj} - C^i_{\ jk}. \quad (17.3.14)$$

Next, by Proposition 16.8.3, the local version of the Koszul calculus is based upon the following formula. For a local frame ϵ , we have

$$\nabla e_j = \Gamma^k_{\ ij}\alpha^i \otimes e_k, \quad \nabla_{e_i} e_j = \Gamma^k_{\ ij}e_k. \quad (17.3.15)$$

Next, acting with ∇_{e_i} on the pairing $\alpha^j(e_k) = \delta^j_k$ and using that the covariant derivative is a derivation of the tensor algebra, we obtain

$$\nabla_{e_i} \alpha^j = -\Gamma^k_{ik}\alpha^k, \quad \nabla \alpha^j = -\Gamma^j_{ik}\alpha^i \otimes \alpha^k. \quad (17.3.16)$$

Now, decomposing an arbitrary tensor field with respect to ϵ and α and using these formulas together with the properties of the covariant derivative, one can derive a local formula for the covariant derivative of any tensor field:

$$\nabla_{\partial_k} t^{i_1 \dots i_s}_{j_1 \dots j_r} = \partial_k t^{i_1 \dots i_s}_{j_1 \dots j_r} + \sum_l \Gamma^{ij}_{km} t^{i_1 \dots i_l=m \dots i_s}_{j_1 \dots j_r} - \sum_l \Gamma^m_{kji} t^{i_1 \dots i_s}_{j_1 \dots j_l=m \dots j_r}. \quad (17.3.17)$$

In particular, for a vector field X and a 1-form α we obtain

$$\nabla_{e_i} X = (e_i(X^k) + \Gamma^k_{ij}X^j)e_k, \quad (17.3.18)$$

$$\nabla_{e_i} \alpha = (e_i(\alpha_j) - \Gamma^k_{ij}\alpha_k)\alpha^j. \quad (17.3.19)$$

Using (16.8.19), we get $\nabla X = \alpha^i \otimes \nabla_{e_i} X$ and $\nabla \alpha = \alpha^i \otimes \nabla_{e_i} \alpha$. Clearly, the covariant derivative of any tensor field t may also be decomposed in this way,

$$\nabla t = \alpha^i \otimes \nabla_{e_i} t, \quad (17.3.20)$$

in accordance with the fact that $\nabla t \in \Omega^1(M; T_s^r M)$.

Remark 17.3.3. By Remark 16.2.5(2), it is clear that the local representatives of Ω and \mathbf{R} , as well as of Θ and \mathbf{T} , coincide. Thus,

$$\mathbf{R}(e_k, e_k)e_l = \mathbf{R}^i{}_{jkl}e_i, \quad \mathbf{T}(e_j, e_k) = \mathbf{T}^i{}_{jk}e_i. \quad (17.3.21)$$

Remark 17.3.4 (Holonomic frames). Let $\{x^i\}_{i=1}^n$ be local coordinates on M . Then $\{\partial_j\}$ is the local holonomic frame of TM , and $\{dx^i\}$ is the dual coframe of T^*M . It is holonomic since $[\partial_i, \partial_j] = 0$. In such a frame, the above formulas take the following form

$$\mathcal{A} = \Gamma^i{}_{jk} dx^j \otimes E^k{}_l, \quad (17.3.22)$$

$$\mathbf{R}^i{}_{jkl} = \partial_j \Gamma^i{}_{kl} - \partial_k \Gamma^i{}_{jl} + \Gamma^m{}_{kl} \Gamma^i{}_{jm} - \Gamma^m{}_{jl} \Gamma^i{}_{km}, \quad (17.3.23)$$

$$\mathbf{T}^i{}_{jk} = \Gamma_{jk} - \Gamma^i{}_{kj}, \quad (17.3.24)$$

$$\nabla \partial_j = \Gamma^k{}_{ij} dx^i \otimes \partial_k. \quad (17.3.25)$$

The change from one holonomic frame to another is given by the Jacobi matrix of the coordinate transformation. Thus, in this case the transition function is

$$x \mapsto \tau(x) = \left(\frac{\partial x^i}{\partial x'^l} \right) \quad (17.3.26)$$

and the transformation formula (17.3.8) reads

$$\Gamma^l{}_{mn} = \Gamma^i{}_{jk} \frac{\partial x^j}{\partial x'^m} \frac{\partial x^k}{\partial x'^n} \cdot \frac{\partial x'^l}{\partial x^i} + \frac{\partial^2 x^i}{\partial x'^m \partial x'^n} \frac{\partial x'^l}{\partial x^i}. \quad (17.3.27)$$

It remains to analyze equations (17.2.2) and 17.2.3 in local coordinates. Let $\gamma(t)$ be given by $t \mapsto x^i(t)$ and, correspondingly, $X = X^i \partial_i$ and $\dot{\gamma} = \dot{x}^i \partial_i$. Using Proposition 16.8.10(2,3) we calculate:

$$\nabla_{\dot{\gamma}} X = \nabla_{\dot{x}^i \partial_i} (X^j \partial_j) = (\dot{x}^i X^j \Gamma^k{}_{ij} + \partial_t(X^k) \dot{x}^i) \partial_k, \quad (17.3.28)$$

that is, (17.2.2) reads

$$\frac{dX^k}{dt} + \Gamma^k{}_{ij} \dot{x}^i X^j = 0. \quad (17.3.29)$$

This is a system of first order ODEs, which by standard theorems admits unique local solutions depending smoothly on the initial values $(t_0, X(t_0))$. The solution $t \mapsto X(t)$ proves the parallel transport

$$\text{tra}_\gamma^\nabla(t) : T_{\gamma(t_0)} M \rightarrow T_{\gamma(t)} M. \quad (17.3.30)$$

Inserting $X^i = \dot{x}^i$ into (17.3.29), we obtain the local form of the geodesic equation:

$$\frac{d^2 x^k}{dt^2} + \Gamma^k{}_{ij} \dot{x}^i \dot{x}^j = 0. \quad (17.3.31)$$

This is a system of second order ODEs, which admits unique local solutions depending smoothly on the initial conditions $(t_0, x^i(t_0), \dot{x}^i(t_0))$.

Remark 17.3.5. 1. Consider the exponential map of a linear connection $\mathcal{H}M$. Via \exp , any frame $\iota_u^\theta : \mathbb{R}^n \rightarrow T_m M$ at $m \in M$ provides a local chart on $T_m M$:

$$\varphi := \exp \circ \iota_u^\theta : \mathbb{R}^n \rightarrow U_m. \quad (17.3.32)$$

This is a local diffeomorphism from a 0-neighborhood in \mathbb{R}^n onto a neighborhood $U_m \subset M$ of m . Taking $\kappa := \varphi^{-1}$ we obtain a local chart (U, κ) centered at m which is called the *local geodesic chart*. The local coordinates x^i of that chart are called normal coordinates at m . In normal coordinates, any geodesic takes the form $x^i(t) = a^i \cdot t$. Thus, at m , we obviously have $\Gamma^k_{ij} + \Gamma^k_{ji} = 0$. That is, for vanishing torsion, the Christoffel symbols vanish at m .

2. The parallel transport of a tangent vector along a closed curve yields a geometric interpretation of curvature, which is in accordance with the Ambrose-Singer Theorem. We have

$$\Delta X^i = -\frac{1}{2} R^i_{jkl} x^l \cdot f^{jk}, \quad (17.3.33)$$

where f^{jk} is a bivector field characterizing the plane enclosed by γ and Δ is the regular Laplace operator on \mathbb{R}^n .

3. The quantity

$$a^i := \frac{dx^i}{dt^2} + \Gamma^i_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} \quad (17.3.34)$$

is the natural generalization of the notion of acceleration of a point particle to curved space. For $a^i = 0$, the particle follows a geodesic.

17.4 Connections on G -structures

In Section 14.4 we have defined G -structures and compatible connections on a **FB** as bundle reductions of the associated principal bundles. We now specialize this to the case of tangent and frame bundles, which is indeed the most common context for studying G -structures. In this case we talk simply of *G -structures on a manifold*.

Definition 17.4.1 (G -structure on a manifold). Let M be a smooth manifold of dimension n .

1. A G -structure on M is a bundle reduction $P \subset \text{Fr } M$ of the frame bundle to a Lie subgroup $G < \text{GL}_n(\mathbb{R})$.
2. A G -structure P is called *integrable* if for every $m \in M$ there exists a local chart (U, κ) with local coordinates x^i such that the induced holonomic frame $\{\partial_i\}$ is a local section of P . Such local coordinates are called admissible.

3. Let $f : M \rightarrow M$ be a diffeomorphism. If $f_* : TM \rightarrow TM$ leaves P invariant, then f is called an automorphism of the G -structure.

Recall that by Corollary 14.4.6, reductions of $\text{Fr } M$ to $G < \text{GL}_n(\mathbb{R})$ are in one-to-one correspondence with smooth sections of the associated bundle $\text{Fr } M \times_{\text{GL}_n(\mathbb{R})} (\text{GL}_n(\mathbb{R})/G)$, or, equivalently, with elements of $\text{Hom}_{\text{GL}_n(\mathbb{R})}(\text{Fr } M, \text{GL}_n(\mathbb{R})/G)$. Thus, the existence of a G -structure is a topological problem that can be dealt with by applying methods of obstruction theory, as we will see in the future. In particular, if $\text{GL}_n(\mathbb{R})/H$ is contractible, then a G -structure certainly exists. Note that, geometrically, a G -structure should be viewed as a bundle of distinguished frames on M .

Recall that a linear connection on M is called *compatible with*, or *adapted to*, the G -structure P if it is reducible to P . Recall also Proposition 16.7.12 characterizing the reducibility of connections in terms of equivariant maps.

Proposition 17.4.2 ([17, Prop. 2.2.3]). *Let P be a G -structure on M and let $\tilde{\varphi} : \text{Fr } M \rightarrow \text{GL}_n(\mathbb{R})/G$ be the $\text{GL}_n(\mathbb{R})$ -equivariant map defining P . Assume that $\text{GL}_n(\mathbb{R})/G$ is embedded into a $\text{GL}_n(\mathbb{R})$ -module F . Then a linear connection ω on $\text{Fr } M$ is compatible with the G -structure P iff $\tilde{\varphi}$ is parallel with respect to ω :*

$$D^\omega \tilde{\varphi} = 0. \quad (17.4.1)$$

Proof. By the proof of Proposition 14.4.3, $P = \{u \in \text{Fr } M : \tilde{\varphi}(u) = [e]\}$. Thus, the restriction of $D^\omega \tilde{\varphi} = 0$ to P reads

$$\sigma_*(\omega)[e] = 0, \quad (17.4.2)$$

which holds iff ω restricted to P takes values in the Lie algebra of G . Thus is equivalent to it being reducible to P . ■

Clearly, for a given G -structure P we may restrict the soldering form θ of $\text{Fr } M$ to P , and, thus, for any connection ω on P we have a torsion 2-form Θ on P defined by (17.1.8).

Proposition 17.4.3 ([17, Prop. 2.2.4]). *If P is an integrable G -structure on M , then it admits a torsion-free connection.*

Proof. Let $\pi : P \rightarrow M$ be the canonical projection. Let s be an integrable local section of P over $U \subset M$. Taking the tangent bundle of the graph of s and extending it using the right G -action to a distribution on P , we obtain a connection on $\pi^{-1}(U) \subset P$. Then, integrability implies $s^*d\theta = 0$ (since, in the corresponding coordinates x^i , the components of s , computed by θ , are constant) and, thus, vanishing of the torsion. Next, we patch together these local

connections to a connection on P using a **POU**. Since torsion is additive this yields the assertion. ■

Remark 17.4.4. Note that the converse is generally not true. Being torsion-free for a G -structure is equivalent to being “integrable to first order” in a certain strict sense.

Since any other connection ω' on P differs from ω by a horizontal 1-form α on P with values in $\mathfrak{g} = \text{Lie } G$,

$$\Theta' = \Theta + \alpha \wedge \theta, \quad \text{or} \quad (\Theta' - \Theta)(X, Y) = \alpha(X)\theta(Y) - \alpha(Y)\theta(X). \quad (17.4.3)$$

By Remark 17.2.8, Θ and α may be identified with G -equivariant functions

$$\mathcal{T} : P \rightarrow \bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n, \quad \tilde{\alpha} : P \rightarrow (\mathbb{R}^n)^* \otimes \mathfrak{g}, \quad (17.4.4)$$

respectively. Since $G \subset \text{GL}_n(\mathbb{R})$, we have a natural inclusion

$$\iota_{\mathfrak{g}} : \mathfrak{g} \rightarrow \text{End}(\mathbb{R}^n) \cong (\mathbb{R}^n)^* \otimes \mathbb{R}^n. \quad (17.4.5)$$

Thus, under the above identification, $\alpha \wedge \theta$ is a function on P with values in $\bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n$. We claim that it coincides with the image of $\tilde{\alpha}$ under the map

$$\delta : (\mathbb{R}^n)^* \otimes \mathfrak{g} \rightarrow \bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n, \quad \delta := (\mathcal{A} \otimes \text{id}_{\mathbb{R}^n}) \circ (\text{id}_{(\mathbb{R}^n)^*} \otimes \iota_{\mathfrak{g}}), \quad (17.4.6)$$

where $\mathcal{A} : (\mathbb{R}^n)^* \otimes (\mathbb{R}^n)^* \rightarrow \bigwedge^2 (\mathbb{R}^n)^*$ is the anti-symmetrization map. Indeed, using $\tilde{\alpha}(u)(\mathbf{x}) = \alpha(B(\mathbf{x}))$, we calculate

$$(\alpha \wedge \theta)_u(B(\mathbf{x}), B(\mathbf{y})) = (\tilde{\alpha}(u)(\mathbf{x})) \mathbf{y} - (\tilde{\alpha}(u)(\mathbf{y})) \mathbf{x} = (\delta \circ \tilde{\alpha}(u))(\mathbf{x}, \mathbf{y}). \quad (17.4.7)$$

As a result,

$$\mathcal{T}' = \mathcal{T} + \delta(\tilde{\alpha}). \quad (17.4.8)$$

Let

$$\text{pr} : \bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n \rightarrow \text{coker}(\delta) = \left(\bigwedge^2 (\mathbb{R}^n)^* \otimes \mathbb{R}^n \right) / \text{im}(\delta) \quad (17.4.9)$$

be the natural projection. Then the map

$$\tau : P \rightarrow \text{coker}(\delta), \quad \tau(u) := \text{pr}(\mathcal{T}(u)), \quad (17.4.10)$$

does not depend on the choice of the connection. This motivates the following definition.

Definition 17.4.5 (Intrinsic torsion). The mapping τ defined above is called the intrinsic torsion of the G -structure P . Moreover, P is called *torsion-free* if $\tau = 0$.

Clearly τ yields the obstruction to the existence of a torsion-free connection on P .

Proposition 17.4.6 ([17, Prop. 2.2.6]). *Let P be a G -structure. Then*

1. *If ω and ω' are torsion-free connections on P and $\omega' - \omega = \alpha \in \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}}$, then $\tilde{\alpha}(u) \in \ker \delta$ for every $u \in P$. In particular, if $\ker \delta = 0$, then P admits at most one torsion-free connection.*
2. *P has a torsion-free connection iff it is torsion-free.*

Proof. The first assertion follows immediately from (17.4.8). For the second one, if P has a torsion-free connection, then it is clearly torsion-free. We prove the converse: let ω be a connection with torsion Θ . By assumption, $\tau = 0$. Thus $\mathcal{T}(u) \in \text{im } \delta$ for all $u \in P$. That is, there exists an equivariant mapping $\tilde{\alpha} : P \rightarrow (\mathbb{R}^n)^* \otimes \mathfrak{g}$ such that $\mathcal{T} = \delta(\tilde{\alpha})$. Let α be the unique horizontal 1-form on P corresponding to $\tilde{\alpha}$. Then $\omega' = \omega - \alpha$ is torsion-free. ■

In particular, as an immediate consequence, we obtain

Corollary 17.4.7 ([17, Cor. 2.2.7]). *If δ is bijective, then P admits a unique torsion-free connection.*

In summary, denoting by $V = \mathbb{R}^n$ the typical fiber, we have constructed the (obviously) exact sequence of linear maps of vector spaces

$$0 \rightarrow \mathfrak{g}^{(1)} := \ker \delta \rightarrow V^* \otimes \mathfrak{g} \xrightarrow{\delta} \bigwedge^2 V^* \otimes V \xrightarrow{\text{pr}} \text{coker } \delta \rightarrow 0. \quad (17.4.11)$$

The space $\mathfrak{g}^{(1)} = \ker \delta$ is called the first *prolongation* of the subalgebra $\mathfrak{g} < \text{gl}_n(\mathbb{R})$. Since these maps are G -equivariant, given a G -structure P on a manifold M , we get the corresponding exact sequence of morphisms of associated VBs. Namely, we simply apply the association functor $P^{[]}$ to each of the vector spaces above as typical fibers and then follow it with the corresponding bundle isomorphism induced from the soldering map $\vartheta : TM \rightarrow P \times_G \mathbb{R}^n$:

$$0 \rightarrow \ker \Delta \rightarrow T^*M \otimes \text{Ad}(P) \xrightarrow[\text{"Contorsion"}]{\Delta} \bigwedge^2 T^*M \otimes TM \rightarrow \frac{\text{coker } \Delta}{\text{"Torsion"} \text{ "Intrinsic torsion"}} \rightarrow 0, \quad (17.4.12)$$

where $T^*M \otimes \text{Ad}(P)$ is the bundle whose space of sections is the gauge algebra $\text{gau}(P) = \Omega^1(M; \text{Ad}(P))$ (consisting of differences $\tilde{\alpha}$) and Δ computes the corresponding change in the torsion map by $\delta(\tilde{\alpha})$. Note that torsion maps themselves are sections of the bundle

$\wedge^2 T^*M \otimes TM$. Then, by exactness, \mathcal{T} and \mathcal{T}' have the same image in the last bundle, $\text{coker } \Delta$. In other words, any two connections compatible with P have torsions that map to the same section of $\text{coker } \Delta$, and this section is exactly the intrinsic torsion of P , independent of the connection. Finally, sections of $\ker \Delta$ are the contorsions whose addition to the connection form doesn't change the torsion.

Note that the map δ can be written as acting on $\kappa \in \text{Hom}(V, \mathfrak{g})$:

$$\delta : \text{Hom}(V, \mathfrak{g}) \rightarrow \text{Hom}\left(\bigwedge^2 V, V\right), \quad (\delta\kappa)(v, w) = \kappa(v)w - \kappa(w)v, \quad v, w \in V. \quad (17.4.13)$$

This is a special instance of a *Spencer differential*. We may summarize this discussion as follows.

Proposition 17.4.8. *Let P be a G -structure on M , let $\theta : TM \rightarrow P \times_G V$ be the soldering map, and let ω be a principal connection form on P with torsion 2-form $\Theta \in \Omega_{\text{hor}}^2(P; V)^\sigma$. If $\omega' = \omega + \alpha$ is another principal connection on P , then its torsion form is $\Theta' = \Theta + \Delta(\alpha)$, where $\Delta : \Omega_{\text{hor}}^1(P; \mathfrak{g})^{\text{Ad}} \rightarrow \Omega_{\text{hor}}^2(P; V)^\sigma$ is induced from the Spencer differential*

$$\delta : \text{Hom}(V, \mathfrak{g}) \rightarrow \text{Hom}\left(\bigwedge^2 V, V\right), \quad (\delta\kappa)(v, w) = \kappa(v)w - \kappa(w)v, \quad v, w \in V. \quad (17.4.14)$$

Remark 17.4.9. 1. Note that for the above construction to go through, it suffices for \mathfrak{g} to be injectively included into $\mathfrak{gl}_n(\mathbb{R})$, whereas G itself doesn't need to be a subgroup of $\text{GL}_n(\mathbb{R})$. Thus a more general definition of a G -structure on TM is as a λ -reduction of the frame bundle via a homomorphism $\lambda : G \rightarrow \text{GL}_n(\mathbb{R})$ such that $\lambda_* : \mathfrak{g} \rightarrow \mathfrak{gl}_n(\mathbb{R})$ is injective (cf. Section 14.4). This implies that G serves as a covering group of an analytic subgroup of $\text{GL}_n(\mathbb{R})$. The most prominent example of this situation are spin structures.
2. We have already established that the intrinsic torsion is independent of the principal connection ω on P . Moreover, unlike torsion, the intrinsic torsion is independent of the choice of the soldering form θ . Therefore it is truly a characteristic of only P and nothing else. The practical implication of this is that in the structure equation

$$d\theta + \omega \wedge \theta = \Theta = \mathcal{T} \circ (\theta \wedge \theta), \quad (17.4.15)$$

the intrinsic torsion can be identified with the part of \mathcal{T} that cannot be absorbed into $\omega \wedge \theta$ by any choice of one-form $\omega \in \Omega^1(P; \mathfrak{g})$ (pseudo-connection). There is no guarantee of uniqueness because different choices of θ will lead to different torsion \mathcal{T} . Thus the second degree of freedom we have is to deform the soldering form θ so as to absorb as much of \mathcal{T} as possible.

3. For a review of non-Lorentzian G -structures on spacetimes, see [48].

Example 17.4.10 (Orientation). Take $G = \mathrm{GL}_n^+(\mathbb{R})$. Then $\mathrm{GL}_n(\mathbb{R})/G \cong \mathbb{Z}_2$. As we saw in Proposition 10.3.5, a section of the associated bundle $\mathrm{Fr} M \times_{\mathrm{GL}_n(\mathbb{R})} \mathbb{Z}^2$ exists iff M is orientable. In this case, the G -structure consists of those frames which are compatible with a chosen orientation. Note that this G -structure is integrable. Also note that automorphisms of this G -structure are exactly the orientation-preserving diffeomorphisms of M .

Example 17.4.11 (Volume form). Now consider $G = \mathrm{SL}_n(\mathbb{R})$. The defining representation of $\mathrm{GL}_n(\mathbb{R})$ on \mathbb{R}^n induces the following $\mathrm{GL}_n(\mathbb{R})$ -action on $\bigwedge^n(\mathbb{R}^n)^*$:

$$\mathrm{GL}_n(\mathbb{R}) \times \bigwedge^n(\mathbb{R}^n)^* \rightarrow \bigwedge^n(\mathbb{R}^n)^*, \quad (g, v) \mapsto \det(g) \cdot v. \quad (17.4.16)$$

Restricted to $\bigwedge^n(\mathbb{R}^n)^* \setminus \{0\}$, this action is transitive and has the common stabilizer $\mathrm{SL}_n(\mathbb{R})$. Thus,

$$\mathrm{GL}_n(\mathbb{R})/G \cong \bigwedge^n(\mathbb{R}^n)^* \setminus \{0\}. \quad (17.4.17)$$

Via the natural isomorphism $\bigwedge^n T^* M \cong \mathrm{Fr} M \times_{\mathrm{GL}_n(\mathbb{R})} \bigwedge^n(\mathbb{R}^n)^*$, the sections of the associated bundle $\mathrm{Fr} M \times_{\mathrm{GL}_n(\mathbb{R})} (\bigwedge^n(\mathbb{R}^n)^* \setminus \{0\})$ are in one-to-one correspondence with volume forms on M . The $\mathrm{SL}_n(\mathbb{R})$ -structure corresponding to a given volume form v consists of those frames u fulfilling $v = v_0 \circ \bigwedge^n u$, where v_0 is the standard volume form on \mathbb{R}^n . Since $\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R})$ is homotopy equivalent to $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n^+(\mathbb{R})$, M admits an $\mathrm{SL}_n(\mathbb{R})$ -structure iff M is orientable. Moreover, it is easy to show that any $\mathrm{SL}_n(\mathbb{R})$ -structure is integrable. Finally, note that the automorphisms of this structure are the volume-preserving diffeomorphisms of M .

Exercise 17.1. Show that every $\mathrm{SL}_n(\mathbb{R})$ -structure is integrable.

Example 17.4.12 (Pseudo-Riemannian structure). For simplicity let us denote $V = \mathbb{R}^n$. Denote the vector space of symmetric covariant tensors of second rank on V by $S^2 V^*$. Endow V with a pseudo-Euclidean metric $\eta \in S^2 V^*$ with signature (k, l) where $n = k + l$. The defining representation of $\mathrm{GL}(V)$ induces a $\mathrm{GL}_n(\mathbb{R})$ -module structure on $S^2 V$ given by

$$\sigma : \mathrm{GL}(V) \rightarrow \mathrm{Aut}(S^2 V^*), \quad \sigma(g) := (g^{-1})^T \otimes (g^{-1}). \quad (17.4.18)$$

As we have noted before, by the Sylvester Theorem, $\mathrm{GL}_n(\mathbb{R})$ acts transitively on the subspace $S_{(k,l)}^2 V^* \subset S^2 V^*$ consisting of elements with fixed signature, and the stabilizer of η is $\mathrm{O}_{k,l}$, that is,

$$\mathrm{GL}_n(\mathbb{R})/\mathrm{O}_{k,l} \cong S_{(k,l)}^2 V^*. \quad (17.4.19)$$

Thus, $\mathrm{O}_{k,l}$ -structures are in one-to-one correspondence with pseudo-Riemannian metrics \mathbf{g} on M and the $\mathrm{O}_{k,l}$ -structure corresponding to \mathbf{g} coincides with the bundle $\mathrm{Fr}_{\mathrm{O}_{k,l}} M$ of orthonormal frames with respect to \mathbf{g} . If (M, \mathbf{g}) is oriented, then $\mathrm{Fr}_{\mathrm{O}_{k,l}} M$ further reduces to a principal $\mathrm{SO}_{k,l}$ -bundle, denoted by $\mathrm{Fr}_{\mathrm{O}_{k,l}}^+ M$.

Note that $\mathrm{GL}_n(\mathbb{R})/\mathrm{O}_n$ is contractible. Thus, an O_n -structure, that is, a Riemannian metric, always exists. On the contrary, structures with arbitrary signatures (k, l) may not exist. Now we show that the corresponding Spencer differential

$$\delta : V^* \otimes \mathfrak{so}_{k,l} \rightarrow \bigwedge^2 V^* \otimes V \quad (17.4.20)$$

is an isomorphism. Indeed, the dimensions of these two spaces coincide because there is a canonical isomorphism

$$\Phi : \mathfrak{so}_{k,l} \rightarrow \bigwedge^2 V, \quad \Phi(A) = \frac{1}{4} A(\mathbf{e}_i) \wedge \eta^{-1}(\alpha^i), \quad (17.4.21)$$

where \mathbf{e}_i is a \mathbf{g} -orthonormal basis in V and $\{\eta^i\}$ is the dual basis. Denoting $A_{ij} = \mathbf{g}(\mathbf{e}_i, A\mathbf{e}_j)$, we obtain

$$\Phi(A) = \frac{1}{4} \eta^{ij} A(\mathbf{e}_i) \wedge \mathbf{e}_j = \frac{1}{4} A^{ij} \mathbf{e}_j \wedge \mathbf{e}_i. \quad (17.4.22)$$

Now it remains to check that $\ker \delta = 0$. Let $\kappa \in \mathrm{Hom}(V, \mathfrak{so}_{k,l})$. Then, for any $v, w, z \in V$, $\kappa(v) \in \mathfrak{so}_{k,l}$, therefore

$$\eta(\kappa(v)w, z) = -\eta(w, \kappa(v)z) = -\eta(\kappa(v)z, w). \quad (17.4.23)$$

If $\delta\kappa = 0$, then, by definition of κ , we also have $\eta(\kappa(v)w, z) = \eta(\kappa(w)v, z)$. Thus, letting

$$T(v, w, z) := \eta(\kappa(v)w, z), \quad (17.4.24)$$

$\delta\kappa = 0$ becomes equivalent to

$$T(v, w, z) = -T(v, z, w) = T(w, v, z), \quad (17.4.25)$$

i.e. T needs to be symmetric in the first two arguments and anti-symmetric in the last two. This forces $T = 0$. Since η is non-degenerate, this implies that $\kappa = 0$ and $\ker \delta = 0$.

Let us interpret the consequences of this fact. First, since δ is an isomorphism, $\mathrm{coker} \delta = 0$, and so every metric connection (assuming one exists!) can be modified to be torsion-free. Moreover, since $\ker \delta = 0$, by Corollary 17.4.7, any $\mathrm{O}_{k,l}$ -structure admits a *unique* torsion-free connection, called the *Levi-Civita connection*.

Another consequence is that for any G -structure with $G < \mathrm{O}_{k,l}$, we have $\ker \delta = 0$, since δ is the restriction to $\mathbf{g} < \mathfrak{so}_{k,l}$ of the Spencer differential we considered above. Thus the following sequence is exact

$$0 \rightarrow V^* \otimes \mathbf{g} \xrightarrow{\delta} \bigwedge^2 V^* \otimes V \rightarrow \mathrm{coker} \delta \rightarrow 0 \quad (17.4.26)$$

and

$$\dim \text{coker } \delta = n \cdot \left(\binom{n}{2} - \dim \mathfrak{g} \right). \quad (17.4.27)$$

This is the rank of the bundle of metric intrinsic torsions.

Example 17.4.13 (Conformal structure [17, Example 2.2.17]). For $n \geq 3$, consider the conformal group

$$\text{CO}_n := \{a \in \text{GL}_n(\mathbb{R}) : a^T a = cI_n, c > 0\}. \quad (17.4.28)$$

Clearly, $\text{CO}_n = O_n \times \mathbb{R}_+$. By the previous example, $\text{GL}_n(\mathbb{R})$ acts transitively on the space $S_{k,l}^2 V^*$. Thus it also acts transitively on the set of conformal equivalence classes of elements of $S_{(k,l)}^2 V^*$ defined by the relation $\eta \sim \eta'$ iff $\eta' = c\eta$ for some $c > 0$. Clearly, the stabilizer of an element $[\eta]$ is CO_n . Thus, CO_n -structures are in one-to-one correspondence with conformal equivalence classes $[\mathbf{g}]$ of metrics on M , with the equivalence defined as follows: $\mathbf{g}' \sim \mathbf{g}$ iff $\mathbf{g}'_m = e^{\varphi(m)} \mathbf{g}_m$ for all $m \in M$ for some $\varphi \in C^\infty(M)$. The CO_n -structure corresponding to class $[\mathbf{g}]$ is denoted by $\text{Fr}_{\text{CO}_n} M$ and is referred to as the bundle of conformal frames. Since $\text{CO}_n = O_n \times \mathbb{R}_+$, the representation theory of CO_n is essentially obtained as an extension of that of O_n . The irreducible representations of \mathbb{R}_+ on \mathbb{R} are labeled by real numbers $r \in \mathbb{R}$ and are given by

$$\mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}, \quad (t, x) \mapsto t^r x. \quad (17.4.29)$$

The number r is called the *conformal weight* of the representation under consideration. Let us denote this representation by L^r . Then a typical CO -module is a tensor product of an O_n -module with L^r . Note that, with respect to the conformal structure $[\mathbf{g}]$, the tangent and the cotangent bundles can no longer be identified, because they correspond to the representations containing factors L^r and L^{-r} , respectively. Clearly, at the level of **VBS**s over M , the additional factors L^r correspond to building the tensor product with an associated line bundle characterized by r .

In a close relation to the previous example, one can show that a conformal structure is integrable iff it is locally conformally flat, that is, iff for every point of M there is a neighborhood U on which the metric is $\mathbf{g} = e^\varphi \mathbf{g}_0$, where \mathbf{g}_0 is the flat Euclidean metric and $\varphi \in C^\infty(U)$. If this condition holds globally, one says that (M, \mathbf{g}) is conformally flat, or, equivalently, that $(M, [\mathbf{g}])$ is flat. A diffeomorphism $f : M \rightarrow M$ is an automorphism of a CO_n -structure iff there exists a function $\varphi \in C^\infty(M)$ such that $f^* \mathbf{g} = e^\varphi \mathbf{g}$, where \mathbf{g} is a representative of the structure. The automorphism group of a conformal structure $(M, [\mathbf{g}])$ is called the conformal group of $(M, [\mathbf{g}])$, denoted $C(M, [\mathbf{g}])$. The following classical theorem holds:

Theorem 17.4.14 ([46]). Let (M, g) be a connection n -dimensional Riemannian manifold with $n \geq 3$. Then its conformal group $C(M, [g])$ is a Lie group of dimension at most $\frac{1}{2}(n+1)(n+2)$.

Example 17.4.15 (Metric connection [17, Example 2.2.22]). By Example 17.4.12, pseudo-Riemannian manifolds are in one-to-one correspondence with $O_{k,l}$ -structures. Thus, let (M, g) be a pseudo-Riemannian manifold and let $\text{Fr}_{O_{k,l}} M$ be its $O_{k,l}$ -structure. In terms of the corresponding equivariant map \tilde{g} ,

$$\text{Fr}_{O_{k,l}} M = \{u \in \text{Fr } M : \tilde{g}(u) = \eta, \} \quad (17.4.30)$$

where η is the standard inner product on \mathbb{R}^n with signature (k, l) . By Proposition 17.4.2, a linear principal connection ω on M is compatible with the $O_{k,l}$ -structure iff g is parallel with respect to ω . A linear connection fulfilling this condition is called metric. By (17.4.18), the metricity condition $D\tilde{g} = d\tilde{g} + \sigma_*(\omega)\tilde{g} = 0$ reads

$$d\tilde{g} - \left(\omega^T \otimes I_n + I_n \otimes \omega^T \right) (\tilde{g}) = 0. \quad (17.4.31)$$

More explicitly, decomposing ω with respect to the basis $\{E^j_i\}$ in $\mathfrak{gl}_n(\mathbb{R})$ and \tilde{g} with respect to the basis in $S^2(\mathbb{R}^n)^*$ induced from the standard basis of \mathbb{R}^n , (17.4.31) takes the form

$$d\tilde{g}_{jk} - \tilde{g}_{jl}\omega^l{}_k - \tilde{g}_{kl}\omega^l{}_j = 0. \quad (17.4.32)$$

But, on $\text{Fr}_{O_{k,l}} M$ we have $\tilde{g}_{kl} = \eta_{kl}$, and thus $d\tilde{g}_{kl} = 0$. This shows that ω is metric iff its reduction to $\text{Fr}_{O_{k,l}} M$ fulfills

$$\omega_{jk} + \omega_{kj} = 0, \quad (17.4.33)$$

that is, iff the reduction takes values in the Lie algebra $\mathfrak{so}_{k,l}$ (as expected from general theory). Equivalently, the metricity condition is given by $\nabla g = 0$. Since ∇_X is a derivation of the tensor algebra, the latter is equivalent to

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \quad (17.4.34)$$

for all $X, Y, Z \in \mathfrak{X}(M)$. Using this, we can give another classical derivation of the torsion-free Levi-Civita connection. Combining (17.4.34) with (17.1.25),

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \quad 0 = \nabla_X Y - \nabla_Y X - [X, Y]. \quad (17.4.35)$$

Then, by direct inspection,

$$\begin{aligned} 2g(\nabla_X Y, Z) &= X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) + \\ &\quad + g([X, Y], Z) + g([Z, X], Y) + g([Z, Y], X). \end{aligned} \quad (17.4.36)$$

One easily checks that this defines a torsion-free connection, thus it coincides with the Levi-Civita connection.

Remark 17.4.16. We can derive local formulas for the Levi-Civita connection. In contrast to general linear connections, there are two natural types of local frames here: holonomic frames and orthonormal frames. First let $\mathbf{e} = \{e_i\}$ be an arbitrary local frame. Formulas (17.3.12), (17.3.14), (16.8.19) and (17.3.15)-(17.3.19) apply in any frame. By (17.4.36),

$$2g(\nabla_{e_i} e_j, e_k) = e_i(g_{jk}) + e_j(g_{ik}) - e_k(g_{ij}) + C^l{}_{ij}g_{lk} + C^l{}_{ki}g_{lj} + C^l{}_{kj}g_{li}, \quad (17.4.37)$$

where $g_{ij} = g(e_i, e_j)$. Thus,

$$\begin{aligned} \Gamma^m{}_{ij} &= \frac{1}{2}g^{mk}(e_i(g_{jk}) + e_j(g_{ik}) - e_k(g_{ij})) + \\ &\quad + \frac{1}{2}\left(C^m{}_{ij} + g^{km}g_{lj}C^l{}_{ki} + g^{km}g_{li}C^l{}_{kj}\right). \end{aligned} \quad (17.4.38)$$

For the case of a holonomic frame $\mathbf{e} = \{\partial_i\}$, we have $g_{ij} = g(\partial_i, \partial_j)$ and $C^i{}_{jk} = 0$, so

$$\Gamma^m{}_{ij} = \frac{1}{2}g^{mk}(g_{jk,i} + g_{ki,j} - g_{ji,k}), \quad \Gamma^m{}_{ij} = \Gamma^m{}_{ji}, \quad (17.4.39)$$

where $g_{ij,k} = \partial_k g_{ij}$. For the case of an orthonormal frame $\mathbf{e} = \{e_i\}$, we have $g_{ij} = \eta_{ij}$, so

$$\Gamma^m{}_{ij} = \frac{1}{2}\left(C^m{}_{ij} + \eta^{km}\eta_{lj}C^l{}_{ki} + \eta^{km}\eta_{li}C^l{}_{kj}\right). \quad (17.4.40)$$

Thus, $\Gamma_{kij} = \eta_{km}\Gamma^m{}_{ij} = \frac{1}{2}(C_{kij} + C_{jki} + C_{ikj})$, and, consequently, in this case we have

$$\Gamma_{kij} = -\Gamma_{jik}, \quad \Gamma^k{}_{ik} = 0. \quad (17.4.41)$$

Using (17.3.15), we obtain

$$d\alpha^i(e_j, e_k) = -\alpha^i([e_j, e_k]) = \Gamma^i{}_{kj} - \Gamma^i{}_{jk} \quad (17.4.42)$$

and, thus,

$$d\alpha^i = -\Gamma^i{}_{jk}\alpha^j \wedge \alpha^k. \quad (17.4.43)$$

Comparing with (17.3.15), we read off

$$d\alpha^i = \alpha^j \wedge \nabla_{e_j} \alpha^i. \quad (17.4.44)$$

This implies the following useful formula for any $\xi \in \Omega^k(M)$:

$$d\xi = \alpha^j \wedge \nabla_{e_j} \xi. \quad (17.4.45)$$

Since the operator d is intrinsic, this formula is independent of the choice of the frame. It can be rewritten as

$$d\xi(e_0, \dots, e_k) = \sum_j (-1)^j \left(\nabla_{e_j} \xi \right) (e_0, \dots, \hat{e}_j, \dots, e_k), \quad (17.4.46)$$

where the hat denotes an omitted argument. By the locality of ∇ and the multilinearity of ξ , we conclude

$$d\xi(X_0, \dots, X_k) = \sum_j (-1)^j \left(\nabla_{X_j} \xi \right) (X_0, \dots, \hat{X}_j, \dots, X_k) \quad (17.4.47)$$

for any set of vector fields $X_0, \dots, X_k \in \mathfrak{X}(M)$.

Example 17.4.17 (Absolute parallelism). An absolute parallelism, i.e. a global frame $\{e_i\}_{i=1}^n$, is the same thing as an $\{e\}$ -structure. Since the Lie algebra is trivial, there exists only one connection ∇ compatible with an absolute parallelism, and it is flat. Namely, if a vector $X \in TM$ has components X^i with respect to $\{e_i\}$, then its parallel transport to any other point of M is defined as the vector with the same components. The vector fields e_i are parallel, as well as their dual covector fields α^i :

$$\nabla e_i = \nabla \alpha^j = 0. \quad (17.4.48)$$

Since $\{e\}$ is a subgroup of SO_n , an $\{e\}$ -structure also induces an orientation and a Riemannian metric, defined by $g(e_i, e_j) = \delta_{ij}$. We have $g = \sum_i \alpha^i \otimes \alpha^i$. Thus ∇ is a metric connection. Since $\nabla_{e_i} e_j = 0$, the components of torsion are $T(e_i, e_j) = -[e_i, e_j]$. Thus the torsion tensor is

$$T(X, Y) = -\alpha^i(X)\alpha^j(Y)[e_i, e_j], \quad (17.4.49)$$

and its components in the frame $\{e_i\}$ are

$$T^i{}_{jk} = -C^i{}_{jk}, \quad (17.4.50)$$

where $C^i{}_{jk}$, as usual, are the structure functions of the frame, $[e_j, e_k] = C^i{}_{jke} e_i$.

As a consequence, there are three more natural (but not compatible) connections associated with a global parallelism:

1. The dual connection $\tilde{\nabla}$ given by $\tilde{\nabla}_X Y := \nabla_Y X + [X, Y]$,
2. The symmetric connection $\hat{\nabla}$ given by $\hat{\nabla}_X Y := \frac{1}{2} (\nabla_X Y + \nabla_Y Z + [X, Y])$,
3. The Levi-Civita connection $\mathring{\nabla}$ of the canonical metric g .

It is then easy to see that

1. $\tilde{\nabla}_X Y = \nabla_X Y - T(X, Y)$,
2. $\hat{\nabla}_X Y = \nabla_X Y - \frac{1}{2} T(X, Y) = \frac{1}{2} (\nabla_X Y + \tilde{\nabla}_X Y)$,

3. $\hat{\nabla}$ and $\mathring{\nabla}$ are torsion-free whereas ∇ and $\tilde{\nabla}$ have opposite torsions.

In particular, the Christoffel symbols of these four connections in the frame $\{e_i\}$ are

1. $\Gamma^i_{jk} = 0$,
2. $\tilde{\Gamma}^i_{jk} = C^i_{jk}$,
3. $\hat{\Gamma}^i_{jk} = \frac{1}{2}C^i_{jk}$,
4. $\mathring{\Gamma}^i_{jk} = \frac{1}{2}(C^i_{jk} + C^j_{ik} + C^k_{ij})$ (by (17.4.38)).

Example 17.4.18 (Lie group structure). Suppose M is a parallelizable connected n -dimensional manifold with an absolute parallelism consisting of complete vector fields $\{e_i\}$ whose torsion tensor is covariantly constant, i.e. its components w.r.t. $\{e_i\}$ are just a set of constants

$$\mathsf{T}^i_{jk} = C^i_{jk} = c^i_{jk} = \text{const.} \quad (17.4.51)$$

Now consider the dual frame of the parallelism consisting of 1-forms $\{\alpha^i\}$. Their exterior derivatives are easily computed using the general formula from Proposition 10.2.10. We do it in the same global frame:

$$d\alpha^i(e_j, e_k) = e_j(\delta^i_k) + e_k(\delta^i_j) - \alpha^i([e_j, e_k]) = -c^i_{jk}. \quad (17.4.52)$$

By inserting arbitrary vector fields $X = X^j e_j$, $Y = Y^k e_k$, we get

$$d\alpha^i(X, Y) = -c^i_{jk}\alpha^j(X)\alpha^k(Y). \quad (17.4.53)$$

Since $c^i_{jk} = -c^i_{kj}$, the right hand side can be written as a wedge product $(c^i_{jk}\alpha^j \wedge \alpha^k)(X, Y)$. Let \mathfrak{g} be the n -dimensional Lie subalgebra of $\mathfrak{X}(M)$ spanned by the parallelism $\{e_i\}$, so that its structure constants are c^i_{jk} . Then, introducing the \mathfrak{g} -valued form $\alpha = (\alpha^1, \dots, \alpha^n)$ and the above equation reads

$$d\alpha + \frac{1}{2}[\alpha, \alpha] = 0, \quad (17.4.54)$$

where $\frac{1}{2}[\alpha, \alpha](X, Y) = [\alpha(X), \alpha(Y)]$ as usual, using the Lie bracket in \mathfrak{g} . Since the covariant exterior derivative coincides with d here, this equation is nothing but the structure equation for the torsion, identifying α with the soldering form.

Clearly $\alpha_m : T_m M \rightarrow \mathbb{R}^n \cong \mathfrak{g}$ is an isomorphism at every $m \in M$. Thus $\alpha \in \Omega^1(M; \mathfrak{g})$ satisfies the conditions of Theorem 11.3.16, and M must be a quotient of a connected n -dimensional Lie group G by a discrete subgroup. α plays the role of the Maurer-Cartan form and $\{e_i\}$ are the left-invariant vector fields. The canonical connection of the parallelism then agrees with the canonical connection ∇^- on the Lie group (Example 17.2.18). Thus, Lie groups can be defined as quotients (by a normal discrete subgroup) of simply connected absolute parallelisms with covariantly constant torsion, plus an arbitrary choice of the identity. As a consequence, every such structure is automatically real analytic.

In particular, a torsion-free absolute parallelism is equivalent to a structure of an affine space.

Example 17.4.19 (Almost symplectic structure). Consider $G = \mathrm{Sp}_n(\mathbb{R})$, which is the group of linear transformations of \mathbb{R}^{2n} that preserve the standard symplectic form ω_{2n} , cf. Example 7.3.2(9). There is a one-to-one correspondence between $\mathrm{Sp}_n(\mathbb{R})$ -structures and almost symplectic manifolds, i.e. even-dimensional manifolds M with a non-degenerate 2-form $\omega \in \Omega^2(M)$. The non-degeneracy condition is equivalent to saying that the $2n$ -form $\omega^n = \omega \wedge \cdots \wedge \omega$ is nowhere vanishing. Let $V = \mathbb{R}^{2n}$, so that the Lie algebra is $\mathfrak{sp}(V) \cong S^2 V$. It is easy to check that $\ker \delta \cong S^3 V$ by using ω_{2n} to identify V with V^* , so the Spencer sequence reads

$$0 \rightarrow S^3 V \rightarrow V^* \otimes S^2 V \rightarrow \bigwedge^2 V^* \otimes V \rightarrow \bigwedge^3 V^* \rightarrow 0. \quad (17.4.55)$$

The associated bundle with fiber $\bigwedge^3 V^*$ is simply $\bigwedge^3 T^* M$, so intrinsic torsions are 3-forms on M . The intrinsic torsion is exactly the part of $\nabla \omega$ that doesn't depend on the choice of connection under which ω is parallel. It is also the fully antisymmetric part of $\nabla \omega$ (since it's a 3-form), so it coincides with $d\omega$.

If we want to classify almost symplectic structures based on some natural constraint on the intrinsic torsion, then we need to decompose $\bigwedge^3 V^*$, as a representation of $\mathrm{Sp}_n(\mathbb{R})$, into irreducible representations (if possible). Since we can use the non-degenerate form ω to contract indices, we can use it to reduce a 3-form to a 1-form. This operation may be called “trace”. Thus, 3-forms that have the form $a\omega_{2n} \wedge \alpha$ with some $a \in \mathbb{R}$ and some $\alpha \in V^*$ may be called “diagonal” (the space of these forms is isomorphic to V^* as a representation of $\mathrm{Sp}_n(\mathbb{R})$), and their complement in $\bigwedge^3 V^*$ consists of traceless forms:

$$\bigwedge^3 V^* \cong V^* \oplus \left(\bigwedge^3 V^* \right)_{\text{traceless}}. \quad (17.4.56)$$

These two components can be shown to be irreducible. Now it is easy to classify all choices of sub-representations:

1. The trivial representation $\{0\}$ leads to the torsion-free condition $d\omega = 0$. Such structures are called *symplectic manifolds*. By Darboux's theorem, they are integrable – there are always local coordinates in which ω looks like ω_{2n} . Note the drastic difference with the Riemannian case, where canonical coordinates exist only if curvature vanishes.
2. The representation V^* leads to the condition $d\omega = \omega \wedge \alpha$ for some $\alpha \in \Omega^1(M)$. Since $0 = d^2 \omega = \omega \wedge d\alpha$, we also find $d\alpha = 0$ (assuming $n > 1$, otherwise any α works), thus *locally* $\alpha = df$ and $d(e^{-f}\omega) = 0$. Such structures are called *locally conformally symplectic* (l.c.s.) manifolds.
3. The representation $\left(\bigwedge^3 V^* \right)_{\text{traceless}}$ is more complicated. First note that for any $f \in C^\infty(M)$ we can consider the *Hamiltonian vector field* $X_f := \omega^{-1}(df)$, where $\omega^1 :$

$T^*M \rightarrow TM$ is the fiberwise inverse of ω . The condition that the intrinsic torsion takes values in the bundle associated to $\left(\bigwedge^3 V^*\right)_{\text{traceless}}$ is equivalent to $\mathcal{L}_{X_f} \omega^n = 0$ for all $f \in C^\infty$ (this condition is known as Liouville's Theorem). It is easy to see that this is equivalent to the condition $d\omega^{n-1} = 0$. This fact is the basis of so called *symplectic integrators*, which are used in computer simulations where one wants to preserve volume. Such almost symplectic structures are called *balanced*.

4. If the intrinsic torsion is unconstrained and belongs only to the entire space $\bigwedge^3 V^*$, we call such an almost symplectic structure *generic*.

Example 17.4.20 (Canonical connections on a Lie group, part 6). Picking up from Example 17.2.18, it remains to examine the last special canonical connection on G , ∇^0 . We have already established that it is torsion-free. The natural question is whether it is a metric connection for some metric.

First of all, it is obvious that left-invariant metrics on G , i.e. such that $L_g^* g = g$, are in one-to-one correspondence with metrics on \mathfrak{g} . Further, it is easy to see that bi-invariant metrics, i.e. such that, in addition $R_g^* g = g$, are in correspondence with Ad -invariant metrics on \mathfrak{g} . By definition, with respect to every such metric on \mathfrak{g} , every adjoint operator Ad_g , $g \in G$, is orthogonal.

17.5 Invariant principal connections

17.6 Invariant linear connections

17.7 Cartan connections

18 Riemannian Geometry

18.1 Riemannian manifolds

Definition 18.1.1 (Riemannian metric on a VB). Let $E \xrightarrow{\pi} M$ be a real VB. A (Riemannian) bundle metric on E is a covariant tensor field $h \in \Gamma(T_2^0 E)$ such that

1. h is symmetric, i.e. $h \in \Gamma(\Sigma^2 E^*)$;
 2. h is non-degenerate, i.e. when viewed as a morphism $h \in \text{Hom}(E, E^*)$, it is an isomorphism;
 3. h is positive: $h(e, e) \geq 0$ and $h(e, e) = 0$ iff $e = 0$. This in fact implies Property 2.
- For $e, e' \in E_p$, we denote $\langle e, e' \rangle_g = g_p(e, e')$.

Given a bundle metric on E , all tensor bundles E_s^r inherit natural bundle metrics as well.

Proposition 18.1.2. *The space of all bundle metrics on E is convex, i.e. if h_1 and h_2 are two metrics on E , then all tensors $h_t = t \cdot h_1 + (1 - t) \cdot h_2$, $t \in [0, 1]$, are metrics.*

Proof. Only non-degeneracy is not obvious. If h_t is degenerate, then there exists an $e \in E \setminus E_0$ such that $h_t(e, e) = 0$. Then by positivity, $h_1(e, e) = h_2(e, e) = 0$, which leads to a contradiction. ■

Theorem 18.1.3 (Existence of bundle metrics). *Every VB $E \xrightarrow{\pi} M$ admits a bundle metric.*

Proof. Choose a bundle atlas consisting of local trivializations $\chi_\alpha : E|_{U_\alpha} \rightarrow U_\alpha \times \mathbb{R}^k$, and on the trivial bundles $U_\alpha \times \mathbb{R}^k$ define a metric h_{stand} using the standard metric on \mathbb{R}^k . Then $h_\alpha|_{U_\alpha} := \chi_\alpha^* h_{\text{stand}}$ and

$$h := \sum_{\alpha} \xi_{\alpha} \cdot h_{\alpha}, \quad (18.1.1)$$

where ξ_α is a POU on M subordinate to the atlas. This is obviously a metric on E . ■

Corollary 18.1.4. *With the help of a metric, any VB has local orthonormal frames. The local trivializations induced by such frames have transition functions that preserve distances, i.e. the cocycle takes values in the orthogonal group, $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow O(k)$. In other words, the structure group of any real VB can be reduced from $GL_k(\mathbb{R})$ to $O(k)$. Furthermore, if the bundle is orientable iff the structure group can be further reduced to $SO(k)$.*

Proposition 18.1.5. *Let h a bundle metric on $E \xrightarrow{\pi} M$ and let $D < E$ be a subbundle. Define*

$$D^\perp := \{v \in E \mid h(v, e) = 0 \text{ for all } e \in D\} = \ker(\text{pr}_D), \quad (18.1.2)$$

where pr_D is the fiberwise orthogonal projection onto D . Then D^\perp is a subbundle of E .

Proof. Let $\{E_1, \dots, E_k\}$ be a local orthonormal frame such that $D = \text{Span}(E_1, \dots, E_m)$ for some $m \leq k$ (such local frames exist by the standard Gram-Schmidt orthogonalization procedure). Then the orthogonal projection is locally written as

$$\text{pr}_D \left(\sum_{j=1}^k s^j E_j \right) = \sum_{j=1}^m s^j E_j. \quad (18.1.3)$$

This definition is in fact basis independent, i.e. defines a global bundle morphism $\text{pr}_D : E \rightarrow E$. Therefore $D^\perp = \ker(\text{pr}_D)$ and is a subbundle by Theorem 8.1.21. ■

Definition 18.1.6 (Riemannian manifold). A Riemannian manifold is a pair (M, g) , where M is a smooth manifold and $g \in \Gamma_2^0(M)$ is a metric on the tangent bundle TM .

Definition 18.1.7 (Isometry). A smooth map between two Riemannian manifolds $f : (M, g) \rightarrow (M', g')$ is called an isometry if $f^*g' = g$.

Corollary 18.1.8. Let (M^m, g) be a Riemannian manifold and $S \lhd M$ with $\dim S = k$. Define $D = TS$ and consider the normal bundle $NS := D^\perp < TM|_S$. Since this is a subbundle of TM , we can find local orthonormal $(m - k)$ -frames for NS . If $\text{codim } S = 1$, then $NS = \text{Span}(N)$, where N is a locally defined vector field. N can be extended to a global section of $TM|_S$ iff S is orientable, and given an orientation on M , the direction of N induces an orientation on S .

Definition 18.1.9 (Flat Riemannian manifold). A Riemannian manifold (M, g) is called flat if it is locally isometric to $(\mathbb{R}^m, g_{\text{stand}})$.

Theorem 18.1.10. Let $f \in C^\infty(M, N)$ and let g be a metric on N . Then f^*g is a metric on M iff f is an immersion. In particular, one can always induce metrics on submanifolds.

Proof. Exercise. ■

Theorem 18.1.11. For a Riemannian manifold (M, g) , TFAE:

1. (M, g) is flat;
2. Everywhere on M there exist local coordinates such that $g = \delta_{ij} dx^i \otimes dx^j$;
3. Everywhere on M there exist local holonomic orthonormal frames.

Proof. Exercise. ■

Corollary 18.1.12. All one-dimensional Riemannian manifolds are flat.

Definition 18.1.13 (Distances induced by a Riemannian metric). For any (piecewise smooth) path $\gamma : [0, 1] \rightarrow M$ define the length of γ with respect to a metric g as

$$L_g[\gamma] := \int_0^1 \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))} dt \equiv \int_0^1 \|\dot{\gamma}(t)\| dt. \quad (18.1.4)$$

Moreover, we define the distance between two points $p, q \in M$ as

$$d_g(p, q) := \inf\{L_g[\gamma] \mid \gamma : p \rightsquigarrow q\}. \quad (18.1.5)$$

The length functional has an important property called *reparametrization invariance*: if $\varkappa : [a, b] \rightarrow [0, 1]$ is any diffeomorphism, then

$$L_g[\gamma \circ \varkappa] = \int_a^b \|\dot{\gamma}(\varkappa(t))\| \varkappa'(t) dt = \int_0^1 \|\dot{\gamma}(\varkappa)\| d\varkappa = L_g[\gamma]. \quad (18.1.6)$$

Theorem 18.1.14. *The distance function d_g on M defined above turns M into a metric space, and its metric topology coincides with the original one.*

Proof. See [7, Thm. 13.29], [10, Lem. 6.2] or [11, Lem. 1.4.1]. ■

Definition 18.1.15 (Conformal metrics). Let g and \tilde{g} be two metric on a manifold M . They are called conformal iff there exists a positive function $\Omega \in C^\infty(M)$, $\Omega > 0$, such that $\tilde{g} = \Omega \cdot g$.

Theorem 18.1.16 (Nomizu-Ozeki conformal completion theorem). *Any Riemannian metric g on a manifold M is conformal to a complete metric \tilde{g} (i.e. a metric that makes M into a complete metric space). In particular, every manifold supports a complete metric.*

Proof. We claim that there exists a conformal factor $\Omega > 0$ such that all \tilde{g} -bounded sets are precompact.

First let us show that this claim implies the theorem. Let (x_n) be a Cauchy sequence w.r.t. $d_{\tilde{g}}$. Then (x_n) is \tilde{g} -bounded and precompact, i.e. contains a convergent subsequence. But then (x_n) itself must converge to the same limit, proving completeness.

Now we prove the claim. Let f be an exhaustion function and $V_j = f^{-1}((-\infty, j))$ be the family of precompact sets that exhaust M . Let $K_j = \overline{V_j}$ be the corresponding compact sets and define the distances $\delta_j = \text{dist}_g(\partial K_j, \partial K_{j+1})$ between their boundaries. Choosing a POU $\{\chi_j\}$ subordinate to the covering $\{V_j \setminus K_{j-1}\}$, define $\Omega = \sum_j \frac{1}{j \cdot \delta_j} \chi_j$. Now it is easy to check that if a set A has finite diameter in the metric $\tilde{g} = \Omega \cdot g$, i.e. $\text{diam}_{\tilde{g}} A < \infty$, then $f|_A$ is bounded. Namely, if (x_n) is a sequence that is not contained in any compact set,

i.e. such that $\forall j \exists N : \forall n > N x_n \notin K_j$, then $x_n \in K_{k(n)} \setminus K_{k(n)-1}$ with $k(n) \rightarrow \infty$ and $d_{\tilde{g}}(x_1, x_n) \geq \sum_{j=1}^{k(n)} \frac{1}{j} \rightarrow \infty$. The contrapositive of this statement is that every \tilde{g} -bounded sequence is contained in a compact set. This implies that all \tilde{g} -bounded sets are precompact, which makes $d_{\tilde{g}}$ a complete metric. ■

Remark 18.1.17. In physics, particularly AdS/CFT, the space (M, \tilde{g}) is often referred to as “conformal compactification” (this terminology was introduced by Roger Penrose). However, the choice of a metric doesn’t modify the topology of M in any sense, so it can’t make a non-compact manifold compact. Therefore a proper name for this procedure is conformal completion.

Proposition 18.1.18. *If (M, g) is complete and $S \lhd M$, then $(S, g|_S)$ is also complete.*

Proof. Exercise. The idea is to show first that $d_M(p, q) \leq d_S(p, q)$ for any $p, q \in S$, where d_M and d_S are the distance functions induced by g on M and S , respectively. Then any Cauchy sequence in S is also a Cauchy sequence in M and therefore converges. The limit of the sequence must belong to S because S is embedded (i.e. closed as a subset of M). ■

Definition 18.1.19 (Musical isomorphisms). Recall that a Riemannian metric $g \in \Gamma_2^0(M)$ can be viewed as an isomorphism $\flat := g \in \text{Hom}(TM, T^*M)$. Similarly, $\sharp := g^{-1} \in \Gamma_0^2(M)$ is an isomorphism $g^{-1} \in \text{Hom}(T^*M, TM)$. We write

$$\alpha^\sharp = g^{-1}(\alpha), \quad v^\flat = g(v), \quad \alpha \in T^*M, \quad v \in TM. \quad (18.1.7)$$

In local coordinates,

$$(\alpha^\sharp)^i = g^{ij} \alpha_j, \quad (v^\flat)_i = g_{ij} v^j, \quad (18.1.8)$$

where g^{ij} is a conventional notation for the components of g^{-1} (as a matrix it is literally the inverse of g_{ij} in any local coordinates). The mnemonic behind the names is that \sharp “raises indices”, i.e. converts one-forms into vectors, and \flat “lowers indices”.

As a result, on a Riemannian manifold (M, g) there is a *natural isomorphism* $TM \cong T^*M$, which is absent without an additional structure like a metric.

A Riemannian metric on an orientable manifold naturally induces a volume form as follows.

Theorem 18.1.20. *Let (M^m, g) be an orientable Riemannian manifold. Then there exists a unique volume form $\mu_g \in \Omega^m(M)$ such that for any positively oriented local orthonormal frame $\{e_i\}_{i=1}^m$ on M , one has $\mu_g(e_1, \dots, e_m) = 1$.*

Proof. Fix a positively oriented orthonormal frame $\{e_i\}$ and define the dual local orthonormal frame by $\alpha^i = (e_i)^\flat$ (indeed, $\alpha^i(e_j) = g(e_i, e_j) = \delta_{ij}$). Consider the local volume form

$$\mu_g = \alpha^1 \wedge \cdots \wedge \alpha^m. \quad (18.1.9)$$

We check that this differential form is in fact independent of the choice of the orthonormal frame. Indeed, any other positively oriented orthonormal frame is obtained from $\{e_i\}$ by the application of a transition function $\varphi_{\alpha\beta}$ with $\varphi_{\alpha\beta*} : U_{\alpha\beta} \rightarrow \text{SO}(m)$ (and the entire atlas can be chosen this way due to orientability), and then the local volume form defined w.r.t. the new frame is

$$\varphi_{\alpha\beta}^* \alpha^1 \wedge \cdots \wedge \varphi_{\alpha\beta}^* \alpha^m = \varphi_{\alpha\beta}^* \mu_g = \underbrace{\det(\varphi_{\alpha\beta})}_{+1} \cdot \mu_g = \mu_g. \quad (18.1.10)$$

Therefore, the above local definition of μ_g stitches together to a global volume form. ■

Note that the above ‘‘Riemannian’’ volume form is natural in the category of oriented Riemannian manifolds, i.e. it is invariant under orientation-preserving isometries.

In local (positively ordered) coordinates,

$$\mu_g = \sqrt{\det(g_{ij})} dx^1 \wedge \cdots \wedge dx^m. \quad (18.1.11)$$

Definition 18.1.21 (Levi-Civita tensor). The volume form μ_g viewed as a totally antisymmetric tensor $\epsilon = \mu_g \in \Gamma_m^0(M)$ is called the Levi-Civita tensor. In local (positively ordered) coordinates,

$$\epsilon_{i_1 \cdots i_m} = \sqrt{g} \varepsilon_{i_1 \cdots i_m}, \quad \sqrt{g} := \sqrt{\det(g_{ij})}, \quad (18.1.12)$$

where $\varepsilon_{i_1 \cdots i_m} = \pm 1$ is the ordinary Levi-Civita symbol.³⁵

Exercise 18.1. Show that the locally defined quantity $dx^1 \wedge \cdots \wedge dx^m$ is in fact a globally well-defined scalar density of weight +1, and similarly \sqrt{g} is a scalar tensor density of weight -1.

Proposition 18.1.22. Let (M, g) be a Riemannian manifold and $S \lhd M$ an orientable submanifold embedded via $i : S \hookrightarrow M$ with a chosen unit normal vector field N . Then:

1. $\mu_{\tilde{g}} = i^*(i_N \mu_g)$, where $\tilde{g} = i^*g$ is the induced metric on S ;
2. given another vector field $X \in \Gamma(TM|_S)$, we have

$$i^*(i_X \mu_g) = \langle X, N \rangle_g \cdot \mu_{\tilde{g}}. \quad (18.1.13)$$

Proof. 1. Assuming that the result is true, we have $\mu_{\bar{g}}(E_1, \dots, E_{m-1}) = \mu_g(N, E_1, \dots, E_{m-1})$, and this equals ± 1 , but since the direction of N is induced from the orientation on M , this in fact equals $+1$. Therefore this is indeed the right volume form on S .

2. Decompose $X = \underbrace{\langle X, N \rangle N}_{X^\perp} + \underbrace{(X - X^\perp)}_{X^\top}$, then

$$i^*(i_{X^\top} \mu_g) \underbrace{(X_1, \dots, X_{m-1})}_{\in TS} = \mu_g \underbrace{(X^\top, X_1, \dots, X_{m-1})}_{\text{lin. dependent}} = 0. \quad (18.1.14)$$

Thus

$$i^*(i_X \mu_g) = i^*(i_{X^\top} \mu_g) = \langle X, N \rangle i^*(i_N \mu_g) = \langle X, N \rangle \mu_{\bar{g}}. \quad (18.1.15)$$

■

Corollary 18.1.23. 1. Define the divergence of a vector field with respect to a metric g as

$$(\operatorname{div}_g X) \mu_g = \mathcal{L}_X \mu_g. \quad (18.1.16)$$

Then by Cartan's magic formula

$$(\operatorname{div}_g X) \mu_g = \operatorname{di}_X \mu_g, \quad \operatorname{div}_g X = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} X^i). \quad (18.1.17)$$

2. The Gauss theorem holds in the form

$$\int_M (\operatorname{div}_g X) \mu_g = \int_M \operatorname{di}_x \mu_g = \int_{\partial M} i_X \mu_g = \int_{\partial M} \langle X, N \rangle_g \mu_g|_{\partial M}. \quad (18.1.18)$$

Definition 18.1.24. Given a Riemannian manifold (M, g) , define the gradient of a function $f \in C^\infty(M)$ as

$$\operatorname{grad}_g f = (\operatorname{d}f)^\sharp. \quad (18.1.19)$$

Define the Laplace-Beltrami operator on $C^\infty(M)$ as

$$\nabla_g^2 f := \operatorname{div}_g \operatorname{grad}_g f. \quad (18.1.20)$$

Definition 18.1.25 (Killing vector field). A vector field $X \in \mathfrak{X}(M)$ on a Riemannian manifold (M, g) is called a Killing vector field if its flow consists of isometries, i.e. $\mathcal{L}_X g = 0$, which is equivalent to the Killing equation in components:

$$X^k \partial_k g_{ij} + \partial_i X^k g_{jk} + \partial_j X^k g_{ik} = 0. \quad (18.1.21)$$

Corollary 18.1.26. 1. In components, the Laplace-Beltrami operator reads

$$\nabla_g^2 f = \frac{1}{\sqrt{g}} \partial_k (g^{ik} \sqrt{g} \partial_i f); \quad (18.1.22)$$

2. Killing vector fields form a Lie subalgebra of $\mathfrak{X}(M)$, i.e. the Lie bracket of two Killing vector fields is a Killing vector field.

Proof. Exercise. ■

Remark 18.1.27. Notice that traditionally in multivariable calculus the components of vector fields are defined w.r.t. to an orthonormal frame as $X = \sum_i \hat{X}^i \hat{e}_i$, where $\hat{e}_i = \frac{\partial_i}{\|\partial_i\|} = \frac{\partial_i}{\sqrt{g_{ii}}}$, but in differential geometry we use the components defined regardless of the metric, $X = \sum_i X^i \partial_i$. Thus there is a discrepancy between the formulas you find on, say, Wikipedia, with the ones in differential geometry textbooks (like (18.1.17) and (18.1.22)), given by $\hat{X}^i = X^i \cdot \sqrt{g_{ii}}$ (of course, this only makes sense as long as we use orthogonal coordinates, in which the metric is diagonal). In a general framework where coordinates can't always be assumed to be orthogonal, the basis of ∂_i 's is preferred.

Exercise 18.2. Using (18.1.17) and (18.1.22), derive the standard expressions for the divergence and the Laplace-Beltrami operator in polar, cylindrical, and spherical coordinates.

18.2 Properties of geodesics

The general references for this section are [11, 20].

As we have already pointed out, the length functional $L_g[\gamma]$ is reparametrization invariant. One of the “disadvantages” of this is that it has “too many” extrema. For instance, in Euclidean space, the length functional would be minimized (with fixed boundary conditions $\gamma(0) = p, \gamma(1) = q$) by any path whose image is the segment of the straight line going through p and q , and there are infinitely many such paths corresponding to all possible parametrizations.

Because of this inconvenience, we prefer to work with a functional that is not reparametrization invariant.

Definition 18.2.1 (Energy functional). For a path $\gamma : [a, b] \rightarrow M$ on a Riemannian manifold (M, g) , we define its “kinetic energy functional” as

$$E_g[\gamma] := \frac{1}{2} \int_a^b \|\dot{\gamma}(t)\|^2 dt = \frac{1}{2} \int_a^b g(\dot{\gamma}(t), \dot{\gamma}(t)) dt. \quad (18.2.1)$$

Lemma 18.2.2. For any path $\gamma : [a, b] \rightarrow M$,

$$L_g[\gamma]^2 \leq 2(b-a)E_g[\gamma]. \quad (18.2.2)$$

The equality holds iff $\|\dot{\gamma}\| = \text{const}$.

Proof. By the Hölder inequality, $\int_a^b \|\dot{\gamma}(t)\| dt \leq \sqrt{b-a} \left(\int_a^b \|\dot{\gamma}(t)\|^2 dt \right)^{1/2}$. ■

Lemma 18.2.3 (Geodesic equation). *The Euler-Lagrange equations for the energy functional E_g on (M^m, g) read in local coordinates*

$$\ddot{x}^i(t) + \Gamma_{jk}^i(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0, \quad i = 1, \dots, m, \quad (18.2.3)$$

where $\Gamma_{jk}^i = \frac{1}{2}g^{il}(\partial_k g_{jl} + \partial_j g_{kl} - \partial_l g_{jk})$ are the so-called Christoffel symbols. This differential equation is called the geodesic equation.

Proof. The Lagrangian is $\mathcal{L} = \frac{1}{2}g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)$, so $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} - \frac{\partial \mathcal{L}}{\partial x^i} = 0$ becomes

$$g_{ik}\ddot{x}^k + g_{ji}\ddot{x}^j + \partial_l g_{ik}\dot{x}^l\dot{x}^k + \partial_l g_{ji}\dot{x}^l\dot{x}^j - \partial_i g_{jk}\dot{x}^j\dot{x}^k = 0. \quad (18.2.4)$$

Using the symmetry of g and acting by g^{-1} on the whole equation, we can bring it to the necessary form. ■

Definition 18.2.4 (Geodesic). A path $\gamma : [a, b] \rightarrow M$ on a Riemannian manifold (M, g) is called a geodesic if in (any) local coordinates it solves the geodesic equation (above), i.e. if it is a critical point of the energy functional E_g .

It is easy to check that if $\gamma(t)$ is a geodesic, then

$$\frac{d}{dt} \|\dot{\gamma}(t)\|^2 = 0, \quad (18.2.5)$$

which means the parametrization of a geodesic is proportional to its length. In other words, motion along a geodesic happens with a constant speed. This is the natural generalization of the notion of “inertial motion” on a curved space (at least in the Newtonian context).

Lemma 18.2.5. Let (M, g) be a Riemannian manifold, $p \in M$, and $v \in T_p M$. Then there exist an $\varepsilon > 0$ and a unique geodesic $\gamma_v : [0, \varepsilon] \rightarrow M$ with initial conditions $\gamma(0) = p$, $\dot{\gamma}(0) = v$. Moreover, γ depends smoothly on p and v .

Proof. This is just a direct application of the general local solvability theorem for ODEs. ■

Corollary 18.2.6. If $\gamma[0, \varepsilon] \rightarrow M$ is a geodesic, then $\gamma^\lambda(t) := \gamma(\lambda t) : [0, \varepsilon/\lambda] \rightarrow M$ is also a geodesic with $\dot{\gamma}^\lambda(0) = \lambda \dot{\gamma}(0)$. Since γ_v depends smoothly on v and the unit sphere $S = \{v \in T_p M \mid \|v\| = 1\}$ is compact, $\exists \varepsilon_0 > 0$ such that the domain of γ_v for all $v \in S$ contains $[0, \varepsilon_0]$. Then for all $\|v\| \leq \varepsilon_0$, the geodesics γ_v are defined on $[0, 1]$.

Definition 18.2.7 (Exponential mapping). Consider the set $V_p = \{v \in T_p M \mid \gamma_v \text{ exists on } [0, 1]\}$. This set is non-empty and is an open neighborhood of the origin in $T_p M$ by the local solvability of the geodesic equation. We define the exponential mapping of the Riemannian manifold (M, g) as

$$\exp_p : V_p \rightarrow M, \quad v \mapsto \gamma_v(1). \quad (18.2.6)$$

Theorem 18.2.8. \exp_p maps a neighborhood of $0 \in T_p M$ diffeomorphically onto a neighborhood of $p \in M$.

Proof. Compute the differential of \exp_p at zero:

$$\exp_{p*0}(v) = \frac{d}{dt} \gamma_{t \cdot v}(1) \Big|_{t=0} = \frac{d}{dt} \gamma_v(t) \Big|_{t=0} = v, \quad (18.2.7)$$

therefore $\exp_{p*} = \text{id}_{T_p M}$ and by the InMT 3.5.1, \exp_p is a local diffeomorphism. ■

Definition 18.2.9 (Riemann normal coordinates). If $T_p M$ is identified with \mathbb{R}^n by a choice of an orthonormal frame at p , then $\exp_p^{-1} : U_p \rightarrow T_p M \cong \mathbb{R}^n$ defines a local chart around p whose components are called Riemann normal coordinates.

Proposition 18.2.10. In Riemann normal coordinates at $p \in M$, $g_{ij}(p) = \delta_{ij}$ and $\Gamma_{jk}^i(p) = 0$ (moreover, $\partial_k g_{ij}(p) = 0$).

Proof. $g_{ij}(p) = \delta_{ij}$ by construction (because we specified an orthonormal frame to introduce the normal coordinates).

Straight lines in normal coordinates are geodesic because the line tv , $t \in \mathbb{R}$, $v \in \mathbb{R}^n$ becomes under the exponential mapping $\gamma_{tv}(1) = \gamma_v(t)$. Thus, in these coordinates for $x(t) = tv$, one has $\Gamma_{jk}^i(tv)v^j v^k = 0$ since $\ddot{x} = 0$. Since this holds for all v in a neighborhood of the origin, we have $\Gamma_{jk}^i(p) = 0$.

Finally, this implies that $\partial_m g_{kj} + \partial_k g_{mj} - \partial_j g_{km}(p) = 0$ and thus $\partial_k g_{jm}(p) = 0$. ■

Definition 18.2.11 (Riemann polar coordinates). Introducing polar coordinates $(r, \varphi^1, \dots, \varphi^{n-1})$ on \mathbb{R}^n identified locally with a neighborhood of $p \in M$ via the exponential mapping, we have $g_{rr} = 1$, $\partial_r g_{r\varphi} = 0$, and thus $g_{r\varphi} \equiv 0$ on the entire chart

(here φ denotes all polar angles collectively). Then the metric takes the canonical form $g = (dr)^2 + g_{\varphi\varphi}$ on the entire neighborhood, where $g_{\varphi\varphi}$ denotes the angular part of the metric, which itself has to be positive-definite. These are the Riemann polar coordinates.

Corollary 18.2.12. *For $p \in M$ there exists $r_0 > 0$ such that the Riemann normal coordinates exist on the open ball $B_{r_0}(p)$ (in the metric d_g). Moreover, for any $q \in \partial B_{r_0}(p)$ there exists a unique geodesic of shortest length (equal to r_0) from p to q , and in the Riemann polar coordinates it is the straight line.*

Proof. The existence of the ball is clear. Now we need to choose r_0 sufficiently small to satisfy the remaining conditions.

For any curve $c(t)$ from p to q , let t_0 be the earliest time when $c(t_0) \in \partial B_{r_0}(p)$. Then

$$L_g \left[c \Big|_{[0, t_0]} \right] = \int_0^{t_0} (g_{ij} \dot{c}^i \dot{c}^j)^{1/2} dt \geq \int (g_{rr}(c(t)) \dot{r}^2)^{1/2} dt, \quad (18.2.8)$$

where the inequality follows from $g_{r\varphi} = 0$ and the fact that the angular part of the metric $g_{\varphi\varphi}$ is positive-definite. Furthermore, $g_{rr} \equiv 1$, so

$$L_g \left[c \Big|_{[0, t_0]} \right] = \int_0^{t_0} |\dot{r}| dt \geq \int_0^{t_0} \dot{r} dt = r(t_0) = r_0. \quad (18.2.9)$$

Moreover, the equality in this relation holds if and only if $g_{\varphi\varphi}(\dot{\varphi}, \dot{\varphi}) = 0$, i.e. exactly when $c(t)$ is a straight line. This proves the existence and uniqueness of the minimizing geodesic and that it's a straight line in these coordinates. ■

Definition 18.2.13 (Geodesically convex sets). An open subset of a Riemannian manifold is called geodesically convex if any two points p, q in it can be connected by a geodesic of the minimum length $d_g(p, q)$ that is entirely contained in this set.

Corollary 18.2.14. *If (M, g) is compact, there exists $r_0 > 0$ such that any two points $p, q \in M$ with $d_g(p, q) \leq r_0$ can be connected by exactly one geodesic of length $d_g(p, q)$. This geodesic depends smoothly on p and q . In particular, M can be covered by geodesically convex balls of radius r_0 .*

Theorem 18.2.15. *Geodesically convex neighborhoods:*

1. *Always exist (locally at every point);*
2. *Can be chosen contractible;*
3. *Are intersection-invariant, i.e. the intersection of any two geodesically convex sets is geodesically convex.*

- Proof.*
1. We claim that the neighborhood $U_p = \exp_p(B_\varepsilon(0))$ of p constructed above is a convex neighborhood. Indeed, with a bit of work one can check that any two points (not just when one of them is p , which is the case considered before) in it are connected by a unique minimizing geodesic that is contained in U_p .
 2. Obvious because we chose them diffeomorphic to a ball;
 3. Obvious.
-

Definition 18.2.16. A Riemannian manifold (M, g) is called geodesically complete if \exp_p is defined on all of $T_p M$ for every $p \in M$, i.e. if all geodesics extend to arbitrarily long times.

Theorem 18.2.17 (Hopf-Rinow). *TFAE:*

1. (M, g) is geodesically complete;
2. (M, d_g) is a complete metric space;
3. d_g -bounded closed sets in M are compact;
4. There exists at least one $p \in M$ such that \exp_p is defined on all of $T_p M$.

Moreover, if (M, g) is geodesically complete, then any two points $p, q \in M$ are connected by at least one geodesic of length $d_g(p, q)$.

Proof. See [11, Thm. 1.7.1] or [20, Thm. 10.9]. ■

18.3 Riemannian connections

18.4 Morse theory

Use RS1, Shen (review).

18.5 Geometry of Lie groups

Haar integral, geodesics, curvature, symmetric spaces, Morse theory, approximation of the loop space ΩG by broken geodesics, proof that $\pi_2(G) = 0$.

18.6 Symmetric spaces

18.7 Maximal compact subgroups

Theorem (Cartan). A compact group of isometries of a nonpositively curved complete simply-connected Riemannian manifold has a fixed point. – Using this, prove uniqueness of the maximal compact subgroup up to conjugation (this proof will apply only to semisimple G). – Theorem (Cartan). A connected real Lie group G is diffeomorphic (as a manifold) to $K \times \mathbb{R}^n$ where K is a maximal compact subgroup of G .

18.8 Bott periodicity

19 Complex Geometry

19.1 Complex and Hermitian structures

Definition 19.1.1 (Complex structure on a vector space). If V is a real vector space, then a complex structure on V is an endomorphism $J \in GL(V)$ such that

$$J^2 = -\text{id}_V. \quad (19.1.1)$$

With a chosen complex structure, (V, J) gains a natural structure of a complex vector space where the multiplication by the imaginary unit is defined as

$$i \cdot v := J(v). \quad (19.1.2)$$

Conversely, every complex vector space has a unique J . Clearly for a complex structure to exist, V has to be even-dimensional. If $\dim V = 2n$, then we can choose a basis on it of the form $\{e_1, \dots, e_n, ie_1, \dots, ie_n\}$.

The category $\mathbf{Vect}_{\mathbb{C}}$ consists of complex vector spaces and the morphisms are \mathbb{C} -linear maps $f : V \rightarrow V'$, i.e. $f(Jv) = J'f(v)$ or $f(v + iw) = f(v) + if(w)$.

Note that the only eigenvalue of J on V is i .

Definition 19.1.2 (Complexification). Let V be a real vector space. The complexification of V is defined as $V^{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}$, i.e. $\dim_{\mathbb{R}} V^{\mathbb{C}} = 2 \dim_{\mathbb{R}} V$. If V already has a chosen complex structure J , then it induces a complex structure on $V^{\mathbb{C}}$ given by $J(v \otimes z) = J(v) \otimes z$. If not, we define $J(v \otimes z) := v \otimes iz$. In any case, J has only two eigenvalues $\pm i$.

Thus $V \mapsto V^{\mathbb{C}}$ is a covariant functor $\mathbf{Vect}_{\mathbb{R}} \rightarrow \mathbf{Vect}_{\mathbb{C}}$.

We also define two subspaces $V^{\pm} = \ker(J \mp i \cdot \text{id}_V) = \frac{1}{2}(\text{id} \mp iJ)(V^{\mathbb{C}})$, i.e. the eigenspaces of J corresponding to its two eigenvalues $\pm i$ on $V^{\mathbb{C}}$.

The complexification is naturally isomorphic to $V \oplus V$ as a real vector space. The two components of the sum are the real and the imaginary part: $J(v, w) = (-w, v)$ (assuming V didn't already have a complex structure, otherwise $J(v, w) = (J(w), J(w))$).

Definition 19.1.3 (Realification). The realification of a complex vector space is the forgetful functor $\mathbf{Vect}_{\mathbb{C}} \rightarrow \mathbf{Vect}_{\mathbb{R}}$ given by $(V, J) \mapsto V$. In particular, it preserves the real dimension $\dim_{\mathbb{R}} V^{\mathbb{R}} = \dim_{\mathbb{R}} V$.

Therefore if we first complexify and then realify a vector space, we end up with twice the dimensions, $\dim_{\mathbb{R}} ((V^{\mathbb{C}})^{\mathbb{R}}) = 2 \dim_{\mathbb{R}} V$.

Continue following Huybrechts, Complex Geometry 1.2.

19.2 Complex manifolds

Definition 19.2.1 (Complex structure on a VB). Let $E \xrightarrow{\pi} M$ be a real VB. A *complex structure* on E is a smooth tensor field $J \in \Gamma^1_1(E)$ such that $J^2 = -\text{id}_E$.

Alternatively, $E \xrightarrow{\pi} M$ is a complex bundle if its local trivializations are $\chi_\alpha : E|_{U_\alpha} \rightarrow U_\alpha \times \mathbb{C}^k$ such that the transition functions are \mathbb{C} -linear, i.e. $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \text{GL}_k(\mathbb{R})$. In other words, it is a real vector bundle of rank $2k$ whose structure group was reduced from $\text{GL}_{2k}(\mathbb{R})$ to a subgroup $\text{GL}_k(\mathbb{C}) < \text{GL}_{2k}(\mathbb{R})$ (how exactly this subgroup is embedded depends on the choice of the complex structure).

Definition 19.2.2 (Complex manifold). A complex n -dimensional manifold is a $2n$ -dimensional smooth manifold M with local charts $\varphi_\alpha : U_\alpha \rightarrow \hat{U}_\alpha \subset \mathbb{C}^n$ and transition maps $\varphi_{\alpha\beta} : \varphi_\beta(U_{\alpha\beta}) \rightarrow \varphi_\alpha(U_{\alpha\beta})$ that are *holomorphic*. Morphisms between complex manifolds are holomorphic maps (i.e. smooth maps whose local representatives in the holomorphic charts are holomorphic).

Definition 19.2.3 (Holomorphic VB). A complex bundle $E \xrightarrow{\pi} M$ over a complex manifold M is called holomorphic if its transition functions are holomorphic as maps $\tau_{\alpha\beta} : U_{\alpha\beta} \rightarrow \text{GL}_k(\mathbb{C})$, where $\text{GL}_k(\mathbb{C})$ is viewed as a complex manifold with the standard complex structure (as an open subset of \mathbb{C}^{k^2}). Morphisms between complex bundles are smooth vector bundle maps that are holomorphic in the base and \mathbb{C} -linear in the fibers.

If M is a complex manifold, then TM (the tangent bundle of M as a real manifold) is naturally a complex bundle, because the multiplication by the imaginary unit can be induced from the local trivializations mapping to \mathbb{C}^n .

However, when can a real manifold be turned into a complex one?

Definition 19.2.4 (Almost complex structure). Let M be a real manifold. An almost complex structure on M is a complex structure on TM .

Definition 19.2.5 (Integrable almost complex structure). An almost complex structure J on a real manifold M is called integrable if there is a structure of a complex manifold on M that induces the complex structure J on TM as described above.

Let $(x^1, \dots, x^n, y^1, \dots, y^n)$ be the standard coordinates on \mathbb{C}^n (i.e. $z^j = x^j + iy^j$) representing a local chart on a complex manifold M . We have $TM = \text{Span}\{\partial_{x^j}, \partial_{y^j}\}_{j=1}^n$. Then the induced complex structure on TM reads

$$J(\partial_{x^k}) = \partial_{y^k}, \quad J(\partial_{y^k}) = -\partial_{x^k}. \tag{19.2.1}$$

Definition 19.2.6 (Holomorphic tangent bundle). Let M be a real manifold. Define the complexification of its tangent bundle as $T^{\mathbb{C}}M := TM \otimes_{\mathbb{R}} \mathbb{C}$ (tensor product with the trivial bundle $M \times \mathbb{C}$). If M was already a complex manifold, we extend J to $T^{\mathbb{C}}M$ by \mathbb{C} -linearity. Otherwise we define $J(v \otimes z) = v \otimes iz$.

Furthermore, define the two eigen-subbundles of $T^{\mathbb{C}}$

$$T^{1,0}M := \ker(J - i \cdot \text{id}), \quad T^{0,1}M := \ker(J + i \cdot \text{id}). \quad (19.2.2)$$

If M was a complex manifold and $(x^1, \dots, x^n, y^1, \dots, y^n)$ are coordinates on it as described above, then

$$T^{1,0}M = \text{Span}_{\mathbb{C}} \left\{ \partial_{z^j} \equiv \frac{1}{2}(\partial_{x^j} - i\partial_{y^j}) \right\}, \quad T^{0,1}M = \overline{T^{1,0}M} = \text{Span}_{\mathbb{C}} \left\{ \partial_{\bar{z}^j} \equiv \frac{1}{2}(\partial_{x^j} + i\partial_{y^j}) \right\}. \quad (19.2.3)$$

The subbundles are called the holomorphic and anti-holomorphic tangent bundles, respectively.

Proposition 19.2.7. $T^{\mathbb{C}}M = T^{1,0}M \oplus T^{0,1}M$ and

$$T^{1,0}M = \{v - iJ \cdot v \mid v \in TM\}, \quad T^{0,1}M = \{v + iJ \cdot v \mid v \in TM\}. \quad (19.2.4)$$

Proof. Check that $J(v - iJv) = Jv - iJ^2v = Jv + iv = i(v - iJv)$, therefore the space of vectors $v - iJV$ is the eigenspace of J corresponding to the eigenvalue i , i.e. exactly $T^{1,0}$. ■

Specifying the subbundle $T^{1,0}M < T^{\mathbb{C}}M$ is equivalent to specifying an almost complex structure J .

There is a natural isomorphism $T^{1,0}M \cong TM$ given by $v \in T^{1,0}M \mapsto v^{\mathbb{R}} = v + \bar{v} \in TM \subset T^{\mathbb{C}}M$. Its inverse is $v \in TM \mapsto v^{\mathbb{C}} = \frac{1}{2}(v - iJv)$. In particular,

$$(\partial_{z^j})^{\mathbb{R}} = (\partial_{\bar{z}^j})^{\mathbb{R}} = \partial_{x^k}, \quad (i\partial_{z^j})^{\mathbb{R}} = \partial_{y^k}. \quad (19.2.5)$$

If M is a complex manifold, then $T^{1,0}M$ has a naturally induced structure of a holomorphic bundle because its transition functions are $(\tau_{\alpha\beta})_k^j = \frac{\partial z_\beta^j}{\partial z_\alpha^k}$, which are holomorphic as complex derivatives of holomorphic maps (namely of the transition maps on M).

But when is the converse true, i.e. what conditions on the almost complex structure J are sufficient for it to come from a structure of a complex manifold (i.e. to be integrable)? The following famous theorem answers this question in a manner very similar to the Frobenius theorem.

Theorem 19.2.8 (Newlander-Nirenberg). *Let M be a smooth real manifold and \mathbf{J} and almost complex structure on it. \mathbf{J} is integrable iff the corresponding complex bundle $T^{0,1}M$ is integrable, i.e. $[T^{0,1}M, T^{0,1}M] \subset T^{0,1}M$.*

This is also equivalent to $N_{\mathbf{J}} = 0$, where $N_{\mathbf{J}}$ is the Nijenhuis tensor, which is a tensor of type $(1, 2)$ defined as a section of $\text{Hom}(T^2M, TM)$ by the formula

$$N_{\mathbf{J}}(X, Y) = \underbrace{-\mathbf{J}^2}_{+1}[X, Y] + \mathbf{J}([\mathbf{J}X, Y] + [X, \mathbf{J}Y]) - [\mathbf{J}X, \mathbf{J}Y]. \quad (19.2.6)$$

Proof. See also proof via Cartan's first structure equation in <https://math.berkeley.edu/~bmcmilla/Talks/Newlander-Nirenberg%20Theorem.pdf>.

This is a difficult theorem in analysis. The integrability of $T^{0,1}M$ implies the existence of a convergent Taylor series at every point for all charts of an atlas on M . The theorem is easy to prove if it is already known that M is a real analytic manifold: by Frobenius theorem, $T^{0,1}M$ integrates to local submanifolds of M which by dimension counting are open subsets, therefore the vector fields $\partial_{\bar{z}^j} := \frac{1}{2}(\partial_{x^j} + i\partial_{y^j})$ actually come from holomorphic local coordinates. ■

Remark 19.2.9. In the case $n = 1$ (one-dimensional complex manifolds, a.k.a. Riemann surfaces), one identically has $N_{\mathbf{J}} = 0$ because of extra symmetries, so all almost complex structures on surfaces are integrable.

19.3 Complex differential forms

Since $T^{\mathbb{C}}M = T^{1,0} \oplus T^{0,1}M$, we have

$$(T^{\mathbb{C}}M)^* = (T^{1,0}M)^* \oplus (T^{0,1}M)^* = T_{0,1}^*M \oplus T_{1,0}^*M, \quad (19.3.1)$$

$$T_{0,1}^*M = \text{Span}(dz^j), \quad T_{1,0}^*M = \text{Span}(d\bar{z}^j). \quad (19.3.2)$$

Indeed, $dz^j(\partial_{z^k}) = \delta_k^j$ and $d\bar{z}^j(\partial_{\bar{z}^k}) = \delta_k^j$. Finally,

$$T_{1,0}^*M = \{\alpha \in T_{\mathbb{C}}^{0,1}M \mid \forall v \in T^{0,1}M, \alpha(v) = 0\} = \{\alpha - i\mathbf{J}^*\alpha \mid \alpha \in T^*M\}, \quad (19.3.3)$$

$$T_{0,1}^*M = \{\alpha \in T_{\mathbb{C}}^{1,0}M \mid \forall v \in T^{0,1}M, \alpha(v) = 0\} = \{\alpha + i\mathbf{J}^*\alpha \mid \alpha \in T^*M\}. \quad (19.3.4)$$

We define the corresponding exterior bundles.

Definition 19.3.1. The complex exterior bundles $\wedge^{p,q} M$ are defined by the equalities

$$\bigwedge^k (T_{\mathbb{C}}^* M) = \bigwedge^k (T_{1,0}^* M \oplus T_{0,1}^* M) = \bigoplus_{p+q=k} \bigwedge^p (T_{1,0}^* M) \otimes \bigwedge^q (T_{0,1}^* M) \equiv \bigoplus_{p+q=k} \bigwedge^{p,q} M. \quad (19.3.5)$$

Furthermore, we define the spaces of complex differential forms of degree (p, q) , denoted by $\Omega^{p,q}(M)$, as the sections of these bundles:

$$\Omega^{p,q}(M) = \Gamma\left(\bigwedge^{p,q} M\right). \quad (19.3.6)$$

If $\omega \in \Omega^{p,q}(M)$, then in local coordinates it has the form

$$\omega = \sum_{I,J} \underbrace{\omega_{i_1 \dots i_p j_1 \dots j_q}}_{\omega_{I,J}} \underbrace{dz^i \wedge \dots \wedge dz^{i_p} \wedge d\bar{z}^{j_1} \wedge \dots \wedge d\bar{z}^{j_q}}_{dz^I \wedge d\bar{z}^J}, \quad (19.3.7)$$

where the sum is over all multi-indices, e.g. $I = (i_1, \dots, i_p)$, $1 \leq i_1 < i_2 \dots < i_p \leq n$. Here, the coefficients are $\omega_{I,J} \in C^\infty(U, \mathbb{C})$.

The exterior derivative is extended by \mathbb{C} -linearity:

$$d\omega = \sum_{I,J} d\omega_{I,J} \wedge dz^I \wedge d\bar{z}^J \in \Omega_{\mathbb{C}}^{p+q+1}(M). \quad (19.3.8)$$

Clearly $d\omega$ is uniquely represented as a sum of an element of $\Omega^{p+1,q}(M)$ and an element of $\Omega^{p,q+1}(M)$, i.e.

$$d(\Omega^{p,q}(M)) \subset \Omega^{p+1,q}(M) \oplus \Omega^{p,q+1}(M). \quad (19.3.9)$$

Definition 19.3.2 (Dolbeault operators). The Dolbeault operators $\partial : \Omega^{p,q}(M) \rightarrow \Omega^{p+1,q}(M)$ and $\bar{\partial} : \Omega^{p,q}(M) \rightarrow \Omega^{p,q+1}(M)$ uniquely defined by the formula

$$d\omega = \partial\omega + \bar{\partial}\omega. \quad (19.3.10)$$

Locally, we have

$$\partial\omega = \sum_{k,I,J} \partial_{z^k} \omega_{I,J} dz^k \wedge dz^I \wedge d\bar{z}^J, \quad (19.3.11)$$

$$\bar{\partial}\omega = \sum_{k,I,J} \partial_{\bar{z}^k} \omega_{I,J} dz^k \wedge dz^I \wedge d\bar{z}^J. \quad (19.3.12)$$

These derivatives have the following basic properties:

1. $\partial(\alpha \wedge \beta) = \partial\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge \partial\beta$, and the same for $\bar{\partial}$;
2. $\overline{\partial}\bar{\alpha} = \partial\alpha$;
3. $\partial^2 = \bar{\partial}^2 = \partial\bar{\partial} + \bar{\partial}\partial = 0$.

The last property follows from the decomposition

$$0 = d^2 = (\partial + \bar{\partial})^2 = \underbrace{\partial^2}_{p \rightarrow p+2} + \underbrace{\bar{\partial}^2}_{q \rightarrow q+2} + \underbrace{\partial\bar{\partial} + \bar{\partial}\partial}_{(p,q) \rightarrow (p+1,q+1)} . \quad (19.3.13)$$

Definition 19.3.3 (Holomorphic differential forms). A form $\omega \in \Omega^{p,0}(M)$ is called holomorphic if $\bar{\partial}\omega = 0$, i.e. if all of its components ω_I are holomorphic functions. (For elements of $\Omega^{p,q}(M)$ this property is called “ $\bar{\partial}$ -closed”.)

Definition 19.3.4 (Canonical line bundle). The canonical bundle of a complex manifold M (where $\dim_{\mathbb{C}} M = n$) is defined as $K_M := \wedge^{n,0} M$, i.e. the determinant bundle of the holomorphic cotangent bundle. Its fiber is \mathbb{C} , so it is a complex line bundle. The dual bundle K_M^* is called the anti-canonical bundle (note that $K_M^* \cong \wedge^n(T^{1,0}M)$, not $\wedge^{0,n} M$).

19.4 Maurer-Cartan equation

Huybrechts 6.1 or Voisin. There's a factor of 1/2 in the MC equation missing in huybrechts because of the way he defines wedge products.

20 Symplectic geometry

Part IV

Homology and Cohomology

21 Category Theory II

21.1 Limits and colimits

Recall the notions of products and coproducts, see Section 1.6. We now generalize their construction to define limits and colimits of \mathcal{J} -shaped diagrams.

Definition 21.1.1 (Limits and colimits). Let \mathcal{J} be an index category and F a functor into any category C , $F : \mathcal{J} \rightarrow C$, thus representing a J -shaped diagram in C (see Def. 1.2.6). Then the limit

$$\varprojlim(F) \tag{21.1.1}$$

is defined as the *terminal object in the category of commutative cones* over the diagram $F(\mathcal{J})$, see Fig. 21.1. That is, it is an object $L \in \text{Ob}(C)$ together with a collection of morphisms $L \xrightarrow{\psi_i} F(i)$ for each $i \in \text{Ob}(\mathcal{J})$ such that each triangle

$$\begin{array}{ccc} & L & \\ \psi_i \swarrow & \circlearrowleft & \searrow \psi_j \\ F(i) & \xrightarrow{f_{ij}} & F(j) \end{array} \tag{21.1.2}$$

commutes and for any other object $A \in \text{Ob}(C)$ with a similar collection of morphisms $A \xrightarrow{\phi_i} F(i)$ there exists a unique morphism $A \rightarrow \varprojlim(F)$ such that the entire diagram including $F(\mathcal{J})$, $\{\psi_i\}$, and $\{\phi_i\}$, commutes.

Similarly, the colimit

$$\varinjlim(F) \tag{21.1.3}$$

is an initial co-cone (i.e. with morphisms pointing towards L) over $F(\mathcal{J})$. Like all universal objects, limits and colimits are defined uniquely up to a unique isomorphism (if they exist at all).

Example 21.1.2. 1. $\mathcal{J} = \emptyset$ and F is the empty functor into C . Then $\varprojlim(F)$ is a terminal object in C and $\varinjlim(F)$ is an initial object in C .
 2. $\mathcal{J} = (\bullet \quad \bullet)$ (discrete category with two objects). Then $\varprojlim(F) = F(1) \times F(2)$ and

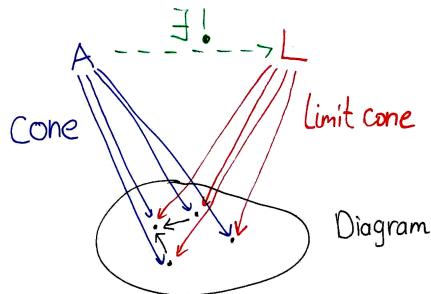


Fig. 21.1: Limit as the terminal cone over a diagram.

- $\lim(F) = F(1) \sqcup F(2).$
- 3. $\overrightarrow{\mathcal{J}} = (A \bullet \rightarrow \bullet_D \leftarrow \bullet_B)$, then $\lim(F) = A \times_D B$.
- 4. $F(\mathcal{J}) = (A \bullet \rightarrow \bullet_B)$, then $\varprojlim(F) = A$ and $\varinjlim(F) = B$.
- 5. $F(\mathcal{J}) = (A \bullet \xrightarrow[\alpha]{\beta} \bullet_B)$. Then $\varprojlim(F) = \text{eq}(\alpha, \beta) = \{a \mid \alpha(a) = \beta(a)\}$ is the equalizer and $\varinjlim(F) = \text{coeq}(\alpha, \beta) = B / (\alpha(a) \sim \beta(a))$ is the coequalizer.

Definition 21.1.3 (Preservation of limits). A covariant functor $G : \mathcal{C} \rightarrow \mathcal{D}$ preserves the limits of the functor $F : \mathcal{J} \rightarrow \mathcal{C}$ if $\lim(G \circ F) = G \lim(F)$. Recall that a limit consists of an object and a set of morphisms from that object that make up the limit cone, so G also acts on these morphisms. Preservation of colimits is defined analogously.

G is said to preserve all J -shaped (co)limits if it preserves the (co)limits of all functors $F : \mathcal{J} \rightarrow \mathcal{C}$.

Note that for contravariant functors the analogous concepts would be functors that take limits to colimits, or colimits to limits.

Definition 21.1.4 (Continuous functors). A functor is called (co)continuous if it preserves all *small* (co)limits (that is, J -shaped (co)limits for all categories J such that $\text{Ob}(J)$ is a set and not a proper class).

Theorem 21.1.5 (Continuity of hom-functors). *Covariant hom-functors $h^A : \mathcal{B} \mapsto \text{Hom}_{\mathcal{C}}(A, B)$ preserve all limits. Similarly, contravariant hom-functors take colimits to limits.*

Remark 21.1.6. Note that the same is not generally true for colimits – e.g. $h^A(B \sqcup C) \neq h^A(B) \sqcup h^A(C)$ in Set .

Corollary 21.1.7. Combining the above theorem with Yoneda's Lemma 1.7.19, we conclude that all representable covariant functors preserve all limits. Similarly, representable contravariant functors map all colimits to limits.

Theorem 21.1.8. Recall that $C^{\mathcal{J}}$ is the category of J-shaped diagrams in C , i.e. covariant functors from \mathcal{J} to C (see Example 1.2.6). Suppose that for each functor $F : \mathcal{J} \rightarrow C$ the limit $\lim_{\leftarrow} F$ exists in C . Then the diagonal functor which assigns to each object A the constant diagram $\underline{A} \in C^{\mathcal{J}}$ (i.e. all objects of \mathcal{J} are mapped to A and all morphisms to id_A),

$$\Delta : C \rightarrow C^{\mathcal{J}}; A \mapsto \underline{A}, \quad (21.1.4)$$

is left adjoint to the limit functor

$$\lim_{\leftarrow} : C^{\mathcal{J}} \rightarrow C; F \mapsto \lim_{\leftarrow} F \quad (21.1.5)$$

(this functor forgets the morphisms associated with the limiting object).

Proof. For the bijection

$$\text{Hom}_{C^{\mathcal{J}}}(\underline{A}, F) \cong \text{Hom}_C(A, \lim_{\leftarrow} F), \quad (21.1.6)$$

note that a natural transformation $\tau : \underline{A} \Rightarrow F$ (which is what morphisms in $C^{\mathcal{J}}$ are) with component morphisms $\tau_j : A \rightarrow F(j)$ (j ranges over all objects in \mathcal{J}) corresponds to a unique morphism $\lim_{\leftarrow} \tau : A \rightarrow \lim_{\leftarrow} F$ in C . Conversely, a morphism $l : A \rightarrow \lim_{\leftarrow} F$ determines a unique natural transformation $\tau : \underline{A} \Rightarrow F$ such that $l = \lim_{\leftarrow} \tau$, namely its components are $\tau_j = l\pi_j$ given the morphisms π_j included in the definition of $\lim_{\leftarrow} F$. It is easy to check that this bijection is natural in A . ■

Corollary 21.1.9. The natural bijection (21.1.6) (with some object L from C instead of $\lim_{\leftarrow} F$ to avoid confusion) holds for each object A in C iff the functor F has a limit.

Proof. The existence of the limit is equivalent to the existence of a unique morphism $\lim_{\leftarrow} \tau : A \rightarrow L$ for each natural transformation $\tau : \underline{A} \Rightarrow F$. ■

Remark 21.1.10. Since a morphism (natural transformation) $\underline{A} \Rightarrow F$ in the category $C^{\mathcal{J}}$ is the same as a cone from A to F , the limit $\lim_{\leftarrow} F$ can be defined as the universal natural transformation $\Delta \Rightarrow F$. Similarly, colimits are the universal natural transformations $F \Rightarrow \Delta$.

Theorem 21.1.11 (Continuity of adjoint functors). *Every functor that has a left adjoint (and therefore is a right adjoint) is continuous. The dual statement is that every functor that has a right adjoint (and therefore is a left adjoint) is cocontinuous.*

Proof. Let $R : C \rightarrow \mathcal{D}$ be the right adjoint to $L : \mathcal{D} \rightarrow C$. Suppose that $\lim_{\leftarrow} F$ exists in C . For each object A in C , a natural bijection

$$\mathrm{Hom}_{C^J}(L(A), F) \cong \mathrm{Hom}_{\mathcal{D}^J}(\underline{A}, R \circ F) \quad (21.1.7)$$

is provided by the natural adjunction bijection

$$\mathrm{Hom}_C(L(A), F(j)) \cong \mathrm{Hom}_{\mathcal{D}}(A, R(F(j))) \quad (21.1.8)$$

for each object j in J , sending the component h_j of an element of the l.h.s. of (21.1.7) to the corresponding component of the natural transformation on the right. Now consider the string of natural bijections

$$\begin{aligned} \mathrm{Hom}_{\mathcal{D}^J}(\underline{A}, R \circ F) &\cong \mathrm{Hom}_{C^J}(L(A), F) \cong \\ &\cong \mathrm{Hom}_C(L(A), \lim_{\leftarrow} F) \cong \mathrm{Hom}_{\mathcal{D}}(A, R(\lim_{\leftarrow} F)) \end{aligned} \quad (21.1.9)$$

coming respectively from (21.1.7), (21.1.6), and the adjunction. By Corollary 21.1.9 combined with the uniqueness of adjoints (Proposition 1.5.13), we have $\lim_{\leftarrow}(R \circ F) = R(\lim_{\leftarrow} F)$ as required. ■

Corollary 21.1.12. *Limits commute with limits (assuming all limits of the necessary shapes exist). Colimits commute with colimits.*

Proof. Viewing the limit as a functor $\lim_{\leftarrow} : C^J \rightarrow C$, we have just shown that it is right adjoint to the diagonal functor (a.k.a. the constant diagram functor). Therefore it is continuous, i.e. commutes with limits. ■

Note that limits and colimits typically do not commute.

Example 21.1.13 (Siefert-van Kampen theorem redux). We've learned by now that contravariant representable (or hom-) functors map colimits to limits, so they're not the functors that preserve colimits. It is the left adjoints that do. Siefert-van Kampen theorem essentially states that the fundamental functors π_1 or Π preserve *some* colimits (namely pushouts). While not a proof, one categorical source of intuition for this is that both of these functors are left adjoints. The reason π_1 preserves only pushouts, and not all colimits, is that the functor that maps the category Top into the homotopy category hTop doesn't preserve colimits, so

π_1 doesn't either. However, the descended version of this functor, $\pi_1 : \text{hTop} \rightarrow \text{Gr}$ happens to be the left adjoint of a certain functor that constructs a topological space given a group. Looking far ahead, this is the functor $B : G \mapsto BG$ that constructs the classifying space of G . It is so called because BG is the base space of a universal principal G -bundle such that any other G -bundle is a pullback along some continuous map into BG . For example, universal covering spaces are exactly the classifying spaces for the fundamental groups of their base – e.g. $B\mathbb{Z} = \mathbb{S}^1$ and $B(\mathbb{Z} * \mathbb{Z}) = \mathbb{S}^1 \vee \mathbb{S}^1$ etc. BG is in fact the unique (up to “weak” homotopy equivalence) space whose fundamental group is G and all other homotopy groups are trivial (which also identifies it with the first Eilenberg-MacLane space $K(G, 1)$). So, by a happy coincidence, pushouts in the homotopy category lined up with pushouts in the original topological category.

The most common examples of infinite limits and colimits are ones where \mathcal{J} can be indexed by integers.

Definition 21.1.14 (Direct and inverse limits). If \mathcal{J} is a directed system and $F : \mathcal{J} \rightarrow C$ a functor, then $\lim_{\rightarrow}(F)$ is called a direct, or an inductive, limit. It is also often written as $\varinjlim A_j$, where $A_j = F(j)$.

If \mathcal{J} is an inverse system (i.e. \mathcal{J}^{op} is a directed system), then $\lim_{\leftarrow}(F)$ is called an inverse, or a projective, limit.

More generally, if \mathcal{J} is a poset category (see Example 1.2.4), then it is said to be directed to the right (left) if $\forall i, j \in \text{Ob}(\mathcal{J}) \exists k \in \text{Ob}(\mathcal{J})$ such that $i, j \leq k$ (respectively, $k \leq i, j$). Then one can define direct limits of diagrams directed to the right and inverse limits of diagrams directed to the left.

Example 21.1.15. 1. In \lim_{\rightarrow} , most commonly all arrows are monomorphisms. For example, consider the category $\overline{\text{Gr}}$ of groups and the sequence S_n of symmetric groups with the embeddings $S_n \hookrightarrow S_m$ for every $n < m$ defined as permutations of the first n symbols. Then $\varinjlim(S_n) = S_{\infty}$, which is the group of all permutations of \mathbb{N} of finite support.

Alternatively, we can define another partial order on the set of symmetric groups S_n . Namely, define $n \preccurlyeq m \Leftrightarrow n \mid m$ and the inclusion $S_n \hookrightarrow S_m$ by m/n copies of the permutation. Then $\lim_{\rightarrow}(S_n)$ is a different group.

2. For any unital ring R , the direct limit of the general linear matrix groups over R (where matrices of size n are embedded into groups of larger matrices by filling in the diagonal with ones) is $\varinjlim(\text{GL}(n, R)) = \text{GL}(R)$, the *quite general linear group* of R .
3. In Set , direct limits are simply infinite unions factored by identifying elements with coincident “descendants”, i.e. $\varinjlim A_i = \bigsqcup_i A_i / \sim$, where $x \in A_i$ is equivalent to $y \in A_j$ iff $\exists f_{jk}, f_{jk}$ such that $f_{ik}(x) = f_{jk}(y)$. The inverse limit is the set of infinite sequences of descendants $\varprojlim(A_i) = \{(a_i) \mid a_i \in A_i, \forall i \leq j, f_{ij}(a_i) = a_j\}$.

4. Consider the polynomial ring $K[t]$ over a field K and its factor rings $K[t]/t^n$ of truncated polynomials. Then we have the sequence of epimorphisms

$$\cdots \rightarrow K[t]/t^3 \rightarrow K[t]/t^2 \rightarrow K, \quad (21.1.10)$$

and the inverse limit is $\varprojlim(K[t]/t^n) = K[[t]]$, which is the ring of *all* formal power series $\sum_{i>0} a_i t^i$. This illustrates the more general fact that projective limits are generally enormous, so much so that the cardinality of the limit is often higher than of any object in the sequence.

5. Let p be a prime integer and consider the sequence of groups

$$\cdots \rightarrow \mathbb{Z}/p^3\mathbb{Z} \xrightarrow{\mod p^2} \mathbb{Z}/p^2\mathbb{Z} \xrightarrow{\mod p} \mathbb{Z}/p\mathbb{Z}. \quad (21.1.11)$$

Then $\varprojlim(\mathbb{Z}/p^n\mathbb{Z}) = \mathbb{Z}_p$ is the group of p -adic integers (it has the cardinality of continuum!).

6. Consider the poset diagram directed to the left consisting of group epimorphisms $\mathbb{Z}/m\mathbb{Z} \xrightarrow{\mod n} \mathbb{Z}/n\mathbb{Z}$ for $n \mid m$. Then $\varprojlim(\mathbb{Z}/n\mathbb{Z}) = \hat{\mathbb{Z}}$ is the *profinite completion of \mathbb{Z}* .
7. The monomorphism sequence $\mathbb{Z}/p^n\mathbb{Z} \xrightarrow{p} \mathbb{Z}/p^{n+1}\mathbb{Z}$ gives the direct limit $\varinjlim(\mathbb{Z}/p^n\mathbb{Z}) = \mathbb{Z}(p^\infty)$, which is called a Prüfer group (it can be realized as the group of all roots of unity of the form $\exp(2\pi i \cdot m/p^n)$).
8. In topological and geometric categories, direct limits are similar to unions (when they exist). For instance, the n -sphere can be embedded $\mathbb{S}^n \hookrightarrow \mathbb{S}^{n+1}$ as the equator, and the direct limit gives $\mathbb{S}^\infty = \varinjlim \mathbb{S}^n$.

21.2 Sub-objects and quotient objects

Definition 21.2.1. Let $X, U, V \in \text{Ob}(C)$ and consider a diagram

$$\begin{array}{ccc} U & & \\ \downarrow \exists h & \nearrow u, \text{mono} & \\ X & & \\ \downarrow v, \text{mono} & \nearrow & \\ V & & \end{array} \quad (21.2.1)$$

We say that $v \leq u$ if there exists a morphism h such that $v = u \circ h$ (h has to be a monomorphism for this to hold).

Here are the properties of this relation:

1. $u \leq u$;
2. $u \leq v, v \leq w \Rightarrow u \leq w$;
3. $u \leq v, v \leq u \Rightarrow U \xrightarrow{f} V$, where $u = v \circ f$ and $v = u \circ f^{-1}$.

Proof. We have $V \xrightarrow{h} U$ and $U \xrightarrow{g} V$.

On the one hand, $v = u \circ h = v \circ g \circ h \Rightarrow g \circ h = 1_V$ since v is mono.

On the other hand, $u = u \circ (h \circ g) \Rightarrow h \circ g = 1_U$ since u is mono. ■

Definition 21.2.2 (Sub-objects). Introduce an equivalence relation on monomorphisms: $u \sim v$ iff $u \leq v$ and $v \leq u$. Then a sub-object of X is an equivalence class of pairs (U, u) , where $u : U \rightarrow X$ is a mono.

Example 21.2.3. $\mathbb{Z} \xrightarrow{\cdot^n} \mathbb{Z}$ is mono and defines the group $n\mathbb{Z}$ as a sub-object of \mathbb{Z} .

Definition 21.2.4 (Quotient objects). For epimorphisms, we define the relation $y \geq z$ if in the diagram

$$\begin{array}{ccc} & Y & \\ y, \text{ epi} \nearrow & \swarrow & \\ X & \curvearrowright & Z \\ z, \text{ epi} \searrow & \nearrow & \\ & \exists h & \end{array} \quad (21.2.2)$$

there exists such a morphism h that $z = h \circ y$ (it has to be epi). Then a quotient object of X is an equivalence class of pairs (Y, y) , where $y : X \rightarrow Y$ is epi.

Example 21.2.5. There are many different epimorphisms $\mathbb{Z}^2 \rightarrow \mathbb{Z}$. They define many non-equivalent quotient objects of \mathbb{Z}^2 , each isomorphic to \mathbb{Z} .

21.3 Abelian categories

Abelian categories are, roughly speaking, categories where objects and morphisms form abelian groups themselves, and where the First Isomorphism theorem holds. Examples include **Ab**, **Vect_K**, **R – Mod** or **Mod – R**, **SheafAb**. Notably, **TopAb** is not abelian because the First Isomorphism theorem generally doesn't hold for topological groups (we call such a category pre-abelian).

Definition 21.3.1 (Abelian categories). A category C is called abelian if:

1. Every morphism set $\text{Mor}(A, B)$ has the structure of an abelian group, i.e. morphisms can be added and subtracted.
2. There exists a zero object $0 \in \text{Ob}(C)$.
3. The product and coproduct of any two objects exist and are isomorphic. The result is denoted by the direct sum: $A \times B \cong A \sqcup B \equiv A \oplus B$.
4. All equalizers and coequalizers exist. In particular, using the abelian property, we define

$$\ker(\varphi) := \text{eq}(\varphi, 0), \quad \text{coker}(\varphi) := \text{coeq}(\varphi, 0). \quad (21.3.1)$$

5. The image and coimage of any morphism coincide, $\text{im}(\varphi) = \text{coim}(\varphi)$. These will be defined below.

We can give alternative definition of kernels and cokernels.

Definition 21.3.2 (Kernel). Let $X \xrightarrow{\varphi} Y$ be a morphism in an abelian category. Then $\ker \varphi$ is the sub-object (K, k) of X such that $\varphi \circ k = 0$ and it is universal with this property:

$$\begin{array}{ccc} K & \xrightarrow{k} & X \xrightarrow{\varphi} Y \\ \exists! h \uparrow \quad \swarrow \psi & & \\ Z & & \end{array} \quad (21.3.2)$$

$\forall \psi: \varphi \circ \psi = 0$

(such a h is automatically unique for every ψ because k is mono).

Definition 21.3.3 (Cokernel). Let $X \xrightarrow{\varphi} Y$ be a morphism in an abelian category. Then $\text{coker } \varphi$ is the quotient object (C, c) of Y such that $c \circ \varphi = 0$ and it is universal with this property:

$$\begin{array}{ccc} X \xrightarrow{\varphi} Y \xrightarrow{c} C \\ \searrow \psi \quad \downarrow \exists! h & & \\ & Z & \end{array} \quad (21.3.3)$$

$\forall \psi: \psi \circ \varphi = 0$

(such a h is automatically unique for every ψ because c is epi).

Example 21.3.4. It is easy to check that in the category \mathbf{Ab} of abelian groups, $\text{coker } \varphi = H/\varphi(G)$ for $\varphi: G \rightarrow H$. Thus the last axiom of abelian categories is equivalent to the First Isomorphism theorem in this case. Note that this factor doesn't exist neither in \mathbf{Gr} nor in \mathbf{TopAb} .

The same formula holds for all ring modules. In \mathbf{Gr} (which is not an abelian category, but in which kernels and cokernels can be similarly defined), the cokernel is the quotient by the

normal closure of the image.

This allows us to define images and coimages.

Definition 21.3.5 (Image, coimage). In abelian categories, $\text{im } \varphi$ for a morphism $\varphi : X \rightarrow Y$ is the sub-object of Y defined as

$$\text{im}(\varphi) = \ker(\text{coker } \varphi), \quad K \xrightarrow{k} X \xrightarrow{\varphi} Y \xrightarrow{c} C. \quad (21.3.4)$$

Similarly, $\text{coim } \varphi$ is the quotient object of X defined as

$$\text{coim}(\varphi) = \text{coker}(\ker \varphi). \quad (21.3.5)$$

Note that

$$\ker(\text{coim } \varphi) = \ker(\text{coker}(\ker \varphi)) = \text{im}(\ker \varphi). \quad (21.3.6)$$

Thus in general pre-abelian categories (i.e. without the last axiom) by the universal properties of images and coimages we have a factorization of any morphism $\varphi : X \rightarrow Y$ into the sequence

$$\ker \varphi \rightarrowtail X \xrightarrow{\text{epi}} \underbrace{\text{coim } \varphi \xrightarrow{\text{bi}} \text{im } \varphi \xrightarrow{\text{mono}} Y}_{\varphi} \twoheadrightarrow \text{coker } \varphi \quad (21.3.7)$$

The last axiom of abelian categories states that $\text{coim } \varphi = \text{im } \varphi$, which is equivalent to the statement:

$$5. \text{ All bimorphisms are isomorphisms.} \quad (21.3.8)$$

Exercise 21.1. Prove that the last axiom of abelian categories is indeed equivalent to the boxed statement.

Therefore only in abelian categories we have the decomposition

$$\ker \varphi \rightarrow X \xrightarrow{j, \text{epi}} \text{im } \varphi \xrightarrow{i, \text{mono}} Y \rightarrow \text{coker } \varphi, \quad (21.3.9)$$

where

$$\text{factorization property : } \varphi = i \circ j, \quad i - \text{mono}, j - \text{epi}. \quad (21.3.10)$$

Definition 21.3.6 (Additive functor). A functor $F : \mathcal{C} \rightarrow \mathcal{D}$ between two abelian categories is called additive if for any $A, B \in \text{Ob}(\mathcal{C})$, the map $F_{A,B} : \text{Hom}_{\mathcal{C}}(A, B) \rightarrow \text{Hom}_{\mathcal{D}}(F(A), F(B))$ defined by $\varphi \mapsto F_{A,B}(\varphi) = F(\varphi)$ is a homomorphism of abelian groups.

The following theorem is the main general result about abelian categories and effectively states that all abelian categories can be realized as (almost arbitrarily nice) full subcategories

of categories of ring modules. It is essentially a much stronger version of the Yoneda Lemma 1.7.19 for abelian categories.

Theorem 21.3.7 (Freyd-Mitchell). *For any abelian category C there exists a ring R and a functor $F : C \rightarrow R\text{-Mod}$ that is: additive, full and faithful (surjective and injective on sets of morphisms for each pair of objects), preserves kernels, cokernels, products and coproducts, is exact (preserves exact sequences, see below)...*

This theorem justifies all diagrammatic methods that we will develop in the next two sections: since $R\text{-Mod}$ is a concrete category, its objects are sets. Therefore by the Freyd-Mitchell theorem, it suffices to prove any general statement about abelian categories only for categories of ring modules, which allows us to refer to elements of objects as sets, and apply morphisms as functions to those elements! It is a way to completely side-step so called “abstract nonsense” proofs that deliberately avoid talking about elements of objects as if they are sets.

22 Homological Algebra I: Exactness

22.1 Exact sequences and functors

From now on in this Part we work only with abelian categories. In fact, homological algebra can be viewed as the extension of linear algebra to all abelian categories.

Definition 22.1.1 (Exact sequences). A sequence of morphisms in an abelian category

$$\cdots \rightarrow A_{i-1} \xrightarrow{\alpha} A_i \xrightarrow{\beta} A_{i+1} \rightarrow \cdots \quad (22.1.1)$$

is called exact in A_i if

$$\text{im } \alpha = \ker \beta. \quad (22.1.2)$$

A sequence is just called exact if it is exact in every object in it.

Proposition 22.1.2. 1. A sequence $0 \rightarrow A \xrightarrow{f} B$ is exact iff f is mono;

2. A sequence $A \xrightarrow{f} B \rightarrow 0$ is exact iff f is epi;

3. A sequence $0 \rightarrow A \xrightarrow{f} B \rightarrow 0$ is exact iff f is an isomorphism;

4. A sequence $0 \rightarrow A \rightarrow B \xrightarrow{f} C \rightarrow D \rightarrow 0$ is exact iff $A = \ker f$ and $D = \text{coker } f$.

Proof. Exercise. ■

Definition 22.1.3 (Short exact sequences). A short exact sequence is an exact sequence of the form

$$0 \rightarrow A \xrightarrow{f} B \xrightarrow{g} C \rightarrow 0. \quad (22.1.3)$$

Such a sequence (and the object B in particular) is also called an *extension* of A by C . The exactness of this sequence is equivalent to f being mono, g being epi, and $\text{im } f = \ker g$.

Proposition 22.1.4 (First isomorphism theorem). If $0 \rightarrow A \xrightarrow{f} B \xrightarrow{g} C \rightarrow 0$ is a short exact sequence, then

$$A \cong \text{im } f, \quad C \cong B/\text{im } f. \quad (22.1.4)$$

Proof. The first isomorphism is known in concrete categories, which is sufficient. It states that $B/\ker g \cong \text{im } g$, and by exactness $\ker g = \text{im } f$ and $\text{im } g = C$, thus $B/\text{im } f \cong C$. ■

Definition 22.1.5 (Split sequence). A short exact sequence $0 \rightarrow A \xrightarrow{i} B \xrightarrow{p} C \rightarrow 0$ is called split if there exists a morphism $j : C \rightarrow B$ with $p \circ j = 1_C$, i.e. if p is a split epi.

Proposition 22.1.6 (Rank-nullity theorem). *If a short exact sequence $0 \rightarrow A \xrightarrow{i} B \xrightarrow{p} C \rightarrow 0$ is split, then $B \cong A \oplus C$.*

Proof. We show that $B \cong \text{im } i \oplus \text{im } j$, where j is a section for p . We perform a simple diagram chasing by considering an element $b \in B$:

$$\begin{aligned} b \in B &\implies p(b - j \circ p(b)) = p(b) - \underbrace{p \circ j}_{1_C}(p(b)) = \\ &= 0 \implies b - j \circ p(b) \in \ker p \stackrel{\text{exact}}{\implies} \exists a \in A : i(a) = b - j \circ p(b). \end{aligned} \quad (22.1.5)$$

This proves that $B = \text{im } i + \text{im } j$. To prove that $\text{im } i \cap \text{im } j = \{0\}$, suppose $x = i(a) = j(c)$. Then $p(x) = p(i(a)) = 0$ since $p \circ i = 0$, and at the same time $p(x) = p(j(c)) = c$ since $p \circ j = 1_C$. Thus $c = 0$, $x = j(c) = 0$, and $B \cong A \oplus C$. ■

Proposition 22.1.7 (Rank-nullity for vector spaces). *If $0 \rightarrow A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} \cdots A_n \rightarrow 0$ is an exact sequence of finite dimensional vector spaces, then*

$$\sum_{i=1}^n (-1)^i \dim A_i = 0. \quad (22.1.6)$$

Proof. By the standard rank-nullity theorem, the l.h.s. equals

$$\sum_{i=1}^{n-1} (-1)^i \dim \ker f_i + \sum_{i=1}^{n-1} (-1)^i \dim \text{im } f_i. \quad (22.1.7)$$

By exactness, this sum vanishes. ■

Definition 22.1.8 (Exact functors). A functor $F : \mathcal{C} \rightarrow \mathcal{D}$ is called exact if it maps every exact sequence into an exact sequence.

Proposition 22.1.9. *For an additive functor F to be exact it suffices to be exact on short sequences.*

Proof. The idea of the proof is to expand any segment of an exact sequence $A \rightarrow B \rightarrow C$ into a combination of short exact sequences (this method is generally called “splicing” of short sequences).

Namely, we have the following *commutative* diagram with exact diagonals:

$$\begin{array}{ccccccc}
 & & & 0 & & & \\
 & & & \searrow & & \swarrow & \\
 & & \ker f & & \text{im } g & & 0 \\
 & & \swarrow & & \nearrow & & \\
 & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \\
 & \downarrow & & \downarrow & & \downarrow & \\
 & \text{im } f & & & & & \text{coker } g \\
 & \downarrow & & \downarrow & & & \downarrow \\
 & 0 & & & 0 & & 0
 \end{array} \tag{22.1.8}$$

Applying F , we get the *commutative* diagram

$$\begin{array}{ccccccc}
 & & & 0 & & & \\
 & & & \searrow & & \swarrow & \\
 & & F(\ker f) & & F(\text{im } g) & & 0 \\
 & & \swarrow & & \nearrow & & \\
 & F(A) & \xrightarrow{F(f)} & F(B) & \xrightarrow{F(g)} & F(C) & \\
 & \downarrow & & \downarrow & & \downarrow & \\
 & F(\text{im } f) & & & & & F(\text{coker } g) \\
 & \downarrow & & \downarrow & & & \downarrow \\
 & 0 & & & 0 & & 0
 \end{array} \tag{22.1.9}$$

with exact diagonals. Now we notice

$$\begin{aligned}
 \text{im } F(f) &= \text{im } (F(A) \rightarrow F(\text{im } f) \rightarrow F(B)) = \text{im } (F(\text{im } f) \rightarrow F(B)) = \\
 &= \ker(F(B) \rightarrow F(\text{im } g)) = \ker(F(B) \rightarrow F(\text{im } g) \rightarrow F(C)) = \ker F(g),
 \end{aligned} \tag{22.1.10}$$

where the second equality follows from $F(A) \rightarrow F(\text{im } f)$ being epi and the next to last equality holds since $F(\text{im } g) \rightarrow F(C)$ is mono. Therefore $F(A) \rightarrow F(B) \rightarrow F(C)$ is exact. ■

There are in fact very few exact functors. Here are a few relevant examples.

Example 22.1.10. 1. Let $G\text{-Ab}$ be the category of abelian groups with a G -action (morphisms in it are equivariant homomorphisms $f(g \cdot a) = g \cdot f(a)$). Then we can define the functor that for every group A with a G -action produces its subgroup of invariants

A^G :

$$A^G = \{a \in A \mid \forall g \in G, g \cdot a = a\}. \quad (22.1.11)$$

(The action on morphisms is trivial.) This is a functor $G\text{-Ab} \rightarrow \text{Ab}$. Now let us consider a short exact sequence

$$0 \rightarrow A \rightarrow B \rightarrow B/A \rightarrow 0. \quad (22.1.12)$$

Its image is clearly not exact on the right:

$$0 \rightarrow A^G \rightarrow B^G \rightarrow (B/A)^G \not\rightarrow 0. \quad (22.1.13)$$

Indeed, $(B/A)^G$ are G -invariant only up to addition of elements of A , whereas B^G/A^G (which is what we would have in a short exact sequence) is a totally different group consisting of classes of truly G -invariant elements. We say that this functor is only *left exact*.

2. The representable (hom-)functors $\text{Hom}_{R\text{-Mod}}(_, _)$ acting $R\text{-Mod} \times R\text{-Mod} \rightarrow \text{Ab}$ with either argument fixed can act on a short exact sequence

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0 \quad (22.1.14)$$

to give two exact sequences

$$0 \rightarrow \text{Hom}(X, A) \rightarrow \text{Hom}(X, B) \rightarrow \text{Hom}(X, C), \quad (22.1.15)$$

$$0 \rightarrow \text{Hom}(C, Y) \rightarrow \text{Hom}(B, Y) \rightarrow \text{Hom}(A, Y). \quad (22.1.16)$$

Therefore Hom-functors are also only left exact. Indeed, right exactness for them would mean that every morphism $X \rightarrow C$ or $A \rightarrow Y$ can be factored through B , which we know to be false. For example, consider

$$\begin{array}{ccccc} 0 & \longrightarrow & A = \mathbb{Z} & \xrightarrow{\text{incl.}} & B = \frac{1}{n}\mathbb{Z} \\ & & \downarrow \text{id} & \swarrow ? & \\ & & Y = \mathbb{Z} & & \end{array} \quad (22.1.17)$$

For right exactness, $\text{id} : A \rightarrow Y$ would need to factor through B , which is clearly impossible here.

An analogous example for the first line is

$$\begin{array}{ccccc} B = \mathbb{Z} & \xrightarrow{\text{mod } n} & C = \mathbb{Z}/n\mathbb{Z} & \longrightarrow & 0 \\ \nwarrow ? & & \uparrow \text{id} & & \\ X = \mathbb{Z}/n\mathbb{Z} & & & & \end{array} \quad (22.1.18)$$

If this functor were to be surjective on Hom-sets, every map $X \rightarrow C$ would have to come from a map $X \rightarrow B$, which is clearly false in this case.

3. The tensor product functor $\underline{\otimes} \underline{} : \text{Mod-}R \times R\text{-Mod} \rightarrow \text{Ab}$ is also not exact. In fact it is only *right exact*: for every short exact sequence

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0 \quad (22.1.19)$$

it gives two exact sequences (in fact they are the same because $X \otimes Y$ is naturally isomorphic to $Y \otimes X$)

$$X \otimes A \rightarrow X \otimes B \rightarrow X \otimes C \rightarrow 0, \quad (22.1.20)$$

$$A \otimes Y \rightarrow B \otimes Y \rightarrow C \otimes Y \rightarrow 0. \quad (22.1.21)$$

We will give an example that shows that this functor is not left exact in Example 22.1.13.

Definition 22.1.11 (Projective and injective objects). If the functor $\text{Hom}(X, \underline{})$ is exact, X is called a projective object (i.e. all morphisms $X \rightarrow C$ factor through B in any exact sequence $B \rightarrow C \rightarrow 0$). If $\text{Hom}(\underline{}, Y)$ is exact, Y is called an injective object (all morphisms $A \rightarrow Y$ factor through B in any exact sequence $0 \rightarrow A \rightarrow B$).

Definition 22.1.12 (Flat module). An object X is called *flat* if the functor $X \otimes \underline{}$ (or equivalently $\underline{} \otimes X$) is left exact.

Example 22.1.13. Consider in the category of \mathbb{Z} -modules the sequence on the left and its image under a tensor product with $\mathbb{Z}/n\mathbb{Z}$

$$0 \rightarrow \mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \xrightarrow{\otimes \mathbb{Z}/n\mathbb{Z}} \mathbb{Z} \otimes \mathbb{Z}/n\mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \otimes \mathbb{Z}/n\mathbb{Z}. \quad (22.1.22)$$

The arrow on the right is obviously not mono (it is the zero morphism), therefore $\mathbb{Z}/n\mathbb{Z}$ is not a flat \mathbb{Z} -module.

The moral of these examples is that whereas in categories of vector spaces Vect_K everything would be exact, exactness is generically broken as soon as we pass to structures over rings. The study of ring modules is a natural extension of linear (matrix) algebra over fields, and largely reduces to homological algebra.

The goal of homological algebra is:
to characterize the obstructions to the exactness of certain additive functors
(these obstructions are called derived functors). (22.1.23)

Given a non-exact functor, the values of its derived functors need to be added into the image of a short exact sequence to produce a fully exact (albeit potentially infinitely long)

sequence. For example, in the case of the invariants functor, every short exact sequence $0 \rightarrow A \rightarrow B \rightarrow B/A \rightarrow 0$ becomes a *long exact sequence of group cohomology*

$$0 \rightarrow A^G \rightarrow B^G \rightarrow (B/A)^G \rightarrow H^1(G, A) \rightarrow H^1(G, B) \rightarrow H^1(G, B/A) \rightarrow H^2(G, A) \rightarrow \dots \quad (22.1.24)$$

For the Hom functor, the obstruction to right exactness is evaluated by the Ext-functors (the name comes from “extension”):

$$0 \rightarrow \text{Hom}(X, A) \rightarrow \text{Hom}(X, B) \rightarrow \text{Hom}(X, C) \rightarrow \text{Ext}^1(X, A) \rightarrow \text{Ext}^1(X, B) \rightarrow \text{Ext}^1(X, C) \rightarrow \dots, \quad (22.1.25)$$

$$0 \rightarrow \text{Hom}(C, Y) \rightarrow \text{Hom}(B, Y) \rightarrow \text{Hom}(A, Y) \rightarrow \text{Ext}^1(C, Y) \rightarrow \text{Ext}^1(B, Y) \rightarrow \text{Ext}^1(A, Y) \rightarrow \dots \quad (22.1.26)$$

Finally, for the tensor functor, the obstruction to left exactness is evaluated by the Tor-functors (“torsion”):

$$\dots \rightarrow \text{Tor}_2(C, X) \rightarrow \text{Tor}_1(A, X) \rightarrow \text{Tor}_1(B, X) \rightarrow \text{Tor}_1(C, X) \rightarrow A \otimes X \rightarrow B \otimes X \rightarrow C \otimes X \rightarrow 0. \quad (22.1.27)$$

All of them are examples of (co)homology theories. We will return to a proper discussion of these objects later. The takeaway so far should be: since on manifolds we are studying spaces of sections of vector bundles, which are really $C^\infty(M)$ -modules, we need homological algebra to deal with the linear algebra over the ring of functions, and the (co)homology groups will measure the non-exactness of certain constructions.

22.2 Direct limits and exactness

Definition 22.2.1. A category C is called complete if all limits in C with a small index category (i.e. $\text{Ob}(\mathcal{J})$ is a proper set and not a class) exist. C is called cocomplete if all small colimits in C exist.

Proposition 22.2.2 ([27, Prop. 5.23]). *Abelian categories are complete and cocomplete.*

Proof. By the Freyd-Mitchell Theorem, it suffices to consider the categories of R -modules. Given a J -shaped diagram $\{F(i) = A_i, f_{ji} = F(i \rightarrow j)\}$ of modules, let λ_i be the natural inclusion of A_i into $\bigoplus_i A_i$.

First we construct the colimit module. Define

$$D = \left(\bigoplus_i A_i \right) / S, \quad (22.2.1)$$

where S is the submodule generated by all elements $\lambda_j \circ f_{ji}(a_i) - \lambda_i(a_i)$ with $a_i \in A_i$ and $i \rightarrow j$ an arrow in \mathcal{J} . The maps λ_i induce inclusions $\alpha_i : A_i \rightarrow D$ by $a_i \mapsto \lambda_i(a_i) + S$. It is an exercise to check that D satisfies the universal property, so that $D \cong \varinjlim F$.

Similarly, the limit module can be constructed as the submodule of $\bigoplus_i A_i$ consisting of tuples $\{a_i\}_i$ such that if $i \rightarrow j$ is an arrow in \mathcal{J} then $a_j = f_{ji}(a_i)$. It is easy to check that this module satisfies the universal property for $\varprojlim F$. ■

Remark 22.2.3. More generally, a category is complete iff all products of any number of objects exist and all equalizers exist, and cocomplete iff all coproducts and all coequalizers exist. In abelian categories both products and coproducts are direct sums and obviously exist, whereas equalizers are $\ker(f - g)$ and coequalizers are $\text{coker}(f - g)$.

In abelian categories, kernels are limits (namely equalizers), and cokernels are colimits. Therefore exactness can be rephrased in terms of preservation of kernels and cokernels. The following Proposition is obvious.

Proposition 22.2.4 ([27, Prop. 5.25]). *Let F be a covariant functor between abelian categories. Then F preserves kernels iff it is left exact, and it preserves cokernels iff it is right exact.*

Proposition 22.2.5. *In an abelian category C , the colimit functor \varinjlim is right exact. That is, if J is small ($\text{Ob}(\mathcal{J})$ is a proper set) and $F, G, H : \mathcal{J} \rightarrow C$ are functors with natural transformations $F \xrightarrow{\alpha} G \xrightarrow{\beta} H$ such that the sequence*

$$0 \rightarrow F(i) \xrightarrow{\alpha_i} G(i) \xrightarrow{\beta_i} H(i) \rightarrow 0 \quad (22.2.2)$$

is exact for all $i \in \text{Ob}(\mathcal{J})$, then the sequence

$$\varinjlim F \varinjlim G \varinjlim H \rightarrow 0 \quad (22.2.3)$$

is exact. Similarly, the limit functor \varprojlim is left exact.

Proof. We make use of the fact that $\varinjlim F$ can be constructed concretely as the quotient of $\bigoplus_i F_i$ by all $a_i - F(i \rightarrow j)(a_i)$ where $i \rightarrow j$ is any arrow coming out of i in \mathcal{J} , $a_i \in F_i$, and $F(i \rightarrow j)(a_i) \in F_j$ are identified with their images in $\bigoplus_i F_i$.

With this description, clearly the map $\varinjlim G(i) \rightarrow \varinjlim H(i)$ is surjective. Also, composing $\varinjlim F(i) \rightarrow \varinjlim G(i) \rightarrow \varinjlim H(i)$ is the zero map thus $\text{im}(\varinjlim \alpha_i) \subset \ker(\varinjlim \beta_i)$.

Conversely, let $x \in \bigoplus_i G(i)$ represent an element of $\ker(\varinjlim \beta)$. Let us define the maps $A = \bigoplus_i \alpha_i$ and $B = \bigoplus_i \beta_i$. Thus $B(x) \in \bigoplus_i H(i)$ is a finite sum of elements of the form $p_i - G(i \rightarrow j)(p_i)$. Since β_i is surjective, we can write such a term as

$$\beta_i(a'_i) - H(i \rightarrow j)(\beta_i(a'_i)) = \beta_i(a'_i) - \beta_j(G(i \rightarrow j)(a'_i)) = B(m'_i - G(i \rightarrow j)(a'_i)) \quad (22.2.4)$$

for some $a'_i \in F(i)$. Since $B(x)$ is a finite sum of $B(a'_i - G(i \rightarrow j))$, we can replace x by another representative such that $B(x) = 0$. Then $x = A(y)$ for some $y \in \bigoplus_i F(i)$. ■

Therefore neither limits nor colimits preserve exactness on both sides, in general. For example pushouts in abelian categories do not preserve exactness. As we will now show, direct limits in abelian categories *are* exact. However, a dual statement for inverse limits does not hold.

Proposition 22.2.6. *Direct limits preserve exactness in abelian categories.*

Proof. It suffices to prove this for direct limits by duality. Right exactness holds for all colimits, so we only need to check left exactness.

Suppose we take a direct limit of modules A_i . Every element of the colimit is represented by some $a_i \in A_i$ for some i . This is because any element of the colimit is represented by some sum of elements $a_j \in A_j$ for various j , and we can pick an index i dominating all of these j 's and take a_i to be the sum of the images of a_j 's in A_i .

Now suppose we have exact sequences $0 \rightarrow A_i \rightarrow B_i \rightarrow C_i \rightarrow 0$ over the same directed system. We want to show that $\lim A_i \rightarrow \lim B_i$ is injective. Pick $a \in \lim A_i$ and let it be represented by $a_i \in A_i$ as above. Now suppose a_i has image 0 in $\lim B_i$. If b_i is the image of s_i , then $b_i = 0$ in the colimit. So for some j , the image of b_i in B_j is 0. So if a_j is the image of a_i in A_j , then $a_j = 0$. Then $a_j = 0$, which makes $A_j \rightarrow B_j = 0$, and so $s_j = 0$ in the colimit. ■

Remark 22.2.7. Elements of inverse limits aren't represented by finite combinations of elements in the A_i 's, therefore the analogous attempt at a proof for inverse limits breaks down. However, inverse limits still sometimes preserve exactness, in particular when the A_i 's satisfy the *Mittag-Leffler condition*: for any $k \in \text{Ob}(\mathcal{J})$ there must exist $i \geq k$ such that for all $j \geq i \geq k$, the two images must coincide, $\text{im}(A_i \rightarrow A_k) \simeq \text{im}(A_j \rightarrow A_k)$. This is satisfied, for example, if all morphisms in the inverse system are surjective.

22.3 Diagram chasing lemmas

All of the following lemmas hold in arbitrary abelian categories. Moreover, for every general diagrammatic statement in an abelian category, its dual holds as well (i.e. the diagram with all arrows reversed), since we can always pass to the dual category, which is also abelian.

Lemma 22.3.1. *If the square*

$$\begin{array}{ccc} A_1 & \xrightarrow{\phi} & B_1 \\ \downarrow \pi & \curvearrowright & \downarrow \rho \\ A_2 & \xrightarrow{\psi} & B_2 \end{array} \tag{22.3.1}$$

commutes, then there exist two morphisms $\eta : \ker \pi \rightarrow \ker \rho$ and $\theta : \text{coker } \pi \rightarrow \text{coker } \rho$ such that

$$\begin{array}{ccc}
\ker \pi & \xrightarrow{\eta} & \ker \rho \\
\downarrow & \circlearrowleft_{\phi} & \downarrow \\
A_1 & \xrightarrow{\quad} & B_1 \\
\downarrow \pi & \circlearrowleft_{\psi} & \downarrow \rho \\
A_2 & \xrightarrow{\quad} & B_2 \\
\downarrow & \circlearrowleft_{\theta} & \downarrow \\
\text{coker } \pi & \xrightarrow{\theta} & \text{coker } \rho
\end{array} \tag{22.3.2}$$

in this diagram the columns, formed by the factorizations of π and ρ , are exact (in fact, exact even after being augmented with zeros on both ends).

Proof. As usual, we only prove this for $R\text{-Mod}$. Let $x \in \ker \pi \subset A_1$. Then $\rho(\phi(x)) = \psi(\pi(x)) = 0$, so $\phi(x) \in \ker \rho$. Thus we define $\eta(x) := \phi(x)$.

For θ , it is in fact enough to pass to the dual category, in which the existence of θ reduces to the above construction of η .

Alternatively, let $x \in A_2/\text{im } \pi$. Define the map

$$A_2/\text{im } \pi \rightarrow B_2/\text{im } \rho, \quad x + \text{im } \pi \mapsto \psi(x) + \text{im } \rho, \tag{22.3.3}$$

which is valid since $\text{im}(\psi \circ \pi) = \text{im}(\rho \circ \phi) \subset \text{im } \rho$.

The commutativity of the resulting diagrams follows from the definitions. ■

Lemma 22.3.2 (3-lemma). *If the rows of the commutative diagram*

$$\begin{array}{ccccc}
A_1 & \xrightarrow{f_1} & B_1 & \xrightarrow{g_1} & C_1 \\
\downarrow \pi & \circlearrowleft_{f_2} & \downarrow \rho & \circlearrowleft_{g_2} & \downarrow \sigma \\
A_2 & \xrightarrow{f_2} & B_2 & \xrightarrow{g_2} & C_2
\end{array} \tag{22.3.4}$$

are exact, then:

1. if σ is mono, then $\text{im } \rho \cap \text{im } f_2 = \text{im}(f_2 \circ \pi) = \text{im}(\rho \circ f_1)$;
2. if π is epi, then $\ker \rho + \text{im } f_1 = \ker(g_2 \circ \rho) = \ker(\sigma \circ g_1)$.

Proof. 1. The inclusion $\text{im}(f_2 \circ \pi) \subset \text{im } \rho \cap \text{im}(f_2)$ is obvious. Now let $x \in \text{im}(\rho) \cap (\text{im}(f_2) = \ker g_2)$. Then $\exists y \in B_1 : \rho(y) = x$. Since $\text{im } f_2 = \ker g_2$, we have

- $g_2(\rho(y)) = 0 = y\sigma(g_1(y))$, so $g_1(y) = 0$ because σ is mono. By the exactness of the top row, $y \in \text{im } f_1$, therefore $\exists z \in A_1 : f_1(z) = y$, thus $x = \rho(f_1(z))$ and $x \in \text{im}(\rho \circ f_1)$.
2. Pass to the dual category and reduce to the first part. Alternatively, the inclusion $\ker \rho + \text{im } f_1 \subset \ker(g_2 \circ \rho)$ is obvious. Now assume $x \in \ker g_2 \circ \rho$. By exactness, $\exists y \in A_2 : f_2(y) = \rho(x)$. Since π is epi, $\exists z \in A_1 : y = \pi(z)$. Then $\rho(f_1(z)) = f_2(\pi(z)) = f_2(y) = \rho(x)$, which means that $f_1(z)$ and x differ by an element of $\ker \rho$, which is what we sought to prove. ■

Lemma 22.3.3 (5-lemma). *If the rows of the commutative diagram*

$$\begin{array}{ccccccc} A_{-2} & \xrightarrow{f_{-2}} & A_{-1} & \xrightarrow{f_{-1}} & A_0 & \xrightarrow{f_0} & A_1 & \xrightarrow{f_1} & A_2 \\ \downarrow \pi_{-2} \curvearrowright & & \downarrow \pi_{-1} \curvearrowright & & \downarrow \pi_0 \curvearrowright & & \downarrow \pi_1 \curvearrowright & & \downarrow \pi_2 \\ B_{-2} & \xrightarrow{g_{-2}} & B_{-1} & \xrightarrow{g_{-1}} & B_0 & \xrightarrow{g_0} & B_1 & \xrightarrow{g_1} & B_2 \end{array} \quad (22.3.5)$$

are exact, then:

1. if π_{-2} is epi and $\pi_{\pm 1}$ are mono, then π_0 is mono;
2. if π_2 is mono and $\pi_{\pm 1}$ are epi, then π_0 is epi.

Proof. Let $x \in A_0$ be such that $\pi_0(x) = 0$. We need to show that $x = 0$. By commutativity, $g_0(\pi_0(x)) = \pi_1(f_0(x)) = 0$, so $f_0(x) = 0$ because π_1 is mono. By exactness, $\exists y \in A_{-1} : f_{-1}(y) = x$, and $g_{-1}(\pi_{-1}(y)) = \pi_0(f_{-1}(y)) = \pi_0(x) = 0$. Then by exactness $\exists z \in B_{-2} : g_{-2}(z) = \pi_{-1}(y)$. Since π_{-2} is epi, $\exists w \in A_{-2} : \pi_{-2}(w) = z$. Now

$$\pi_{-1}(y) = g_{-2}(\pi_{-2}(w)) = \pi_{-1}(f_{-2}(w)) \implies y = f_{-2}(w) \in \text{im } f_{-2} = \ker f_{-1}, \quad (22.3.6)$$

since π_{-1} is mono. Therefore $x = f_{-1}(y) = 0$ by exactness. ■

Corollary 22.3.4. 1. If π_{-2} is epi, π_2 is mono, and $\pi_{\pm 1}$ are iso, then π_0 is iso;
2. If the diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & 0 \\ & & \pi_{-1} \downarrow \curvearrowright & & \pi_0 \downarrow \curvearrowright & & \pi_1 \downarrow & & \\ 0 & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & 0 \end{array} \quad (22.3.7)$$

has exact rows and $\pi_{\pm 1}$ are both epi (mono), then π_0 is epi (respectively, mono).

Example 22.3.5. For any morphism $f : X \rightarrow Y$ we have the commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \ker f & \longrightarrow & X & \longrightarrow & \text{coim } f & \longrightarrow & 0 \\ & & \downarrow 0 & \curvearrowleft & \downarrow f & \curvearrowleft & \downarrow 0 \\ 0 & \longrightarrow & \text{im } f & \longrightarrow & Y & \longrightarrow & \text{coker } f & \longrightarrow & 0 \end{array} \quad (22.3.8)$$

and its rows are exact by the definitions of (co)images and (co)kernels. The two zero morphisms $\ker f \xrightarrow{0} \text{im } f$ and $\text{coim } f \xrightarrow{0} \text{coker } f$ are mono (epi) iff f itself is mono (epi, respectively), which can be seen as an application of the 5-lemma.

Lemma 22.3.6 (4-lemma). *If the rows of the commutative diagram*

$$\begin{array}{ccccccc} A_1 & \longrightarrow & A_2 & \xrightarrow{f} & A_3 & \longrightarrow & A_4 \\ \downarrow \pi_1 & \curvearrowleft & \downarrow \pi_2 & \curvearrowleft & \downarrow \pi_3 & \curvearrowleft & \downarrow \pi_4 \\ B_1 & \longrightarrow & B_2 & \xrightarrow{g} & B_3 & \longrightarrow & B_4 \end{array} \quad (22.3.9)$$

are exact, π_1 is epi, and π_4 is mono, then:

$$\ker \pi_3 = f(\ker \pi_2), \quad \text{im } \pi_2 = g^{-1}(\text{im } \pi_3). \quad (22.3.10)$$

Proof. Exercise. ■

Corollary 22.3.7 (Weak 4-lemma). *In the above diagram, in addition:*

1. if π_3 is epi, then so is π_2 ;
2. if π_2 is mono, then so is π_3 .

Lemma 22.3.8 (Snake lemma/Connecting homomorphism lemma). *Let the rows of the commutative diagram*

$$\begin{array}{ccccccc} A_0 & \xrightarrow{f_0} & A_1 & \xrightarrow{f_1} & A_2 & \longrightarrow & 0 \\ & \downarrow \pi_0 & \curvearrowleft & \downarrow \pi_1 & \curvearrowleft & \downarrow \pi_2 & \\ 0 & \longrightarrow & B_0 & \xrightarrow{g_0} & B_1 & \xrightarrow{g_1} & B_2 \end{array} \quad (22.3.11)$$

be exact. Then with the following diagram

$$\begin{array}{ccccccc}
 \ker \pi_0 & \xrightarrow{\eta_0} & \ker \pi_1 & \xrightarrow{\eta_1} & \ker \pi_2 & \dashrightarrow & \\
 \downarrow & \text{--->} & \downarrow & \text{--->} & \downarrow & & \\
 A_0 & \xrightarrow{f_0} & A_1 & \xrightarrow{f_1} & A_2 & \longrightarrow 0 & \\
 \downarrow \pi_0 & \text{--->} & \downarrow \pi_1 & \text{--->} & \downarrow \pi_2 & & \\
 0 & \xrightarrow{g_0} & B_0 & \xrightarrow{g_1} & B_1 & \xrightarrow{g_2} & B_2 \\
 \downarrow & \text{--->} & \downarrow & \text{--->} & \downarrow & & \\
 \text{coker } \pi_0 & \xrightarrow{\theta_0} & \text{coker } \pi_1 & \xrightarrow{\theta_1} & \text{coker } \pi_2 & & \\
 \dashrightarrow & \text{--->} & \dashrightarrow & \text{--->} & \dashrightarrow & & \\
 \end{array}$$

(22.3.12)

there exists a unique connecting morphism δ shown in the diagram above which makes the kernel-cokernel sequence exact:

$$\ker \pi_0 \rightarrow \ker \pi_1 \rightarrow \ker \pi_2 \xrightarrow{\delta} \text{coker } \pi_0 \rightarrow \text{coker } \pi_1 \rightarrow \text{coker } \pi_2. \quad (22.3.13)$$

Proof. First one checks the exactness of the top and bottom rows of the large diagram using the 3-lemma.

Next we construct δ . Take $x \in \ker \pi_2 \subset A_2$. Then $\exists y \in A_1 : f_1(y) = x$ since f_1 is epi. By commutativity, $\pi_1(y) \in \ker g_1$, and by exactness, $\exists z \in B_0 : g_0(z) = \pi_1(y)$. We define

$$\delta(x) = z + \text{im } \pi_0 = [g_0^{-1} \circ \pi_1 \circ f_1^{-1}(x)] \in B_0/\text{im } \pi_0 = \text{coker } \pi_0. \quad (22.3.14)$$

Note that z is uniquely determined by y because g_0 is mono.

Next we need to check correctness: given another $y' : f_1(y') = x$, we have a unique $z' : g_0(z') = \pi_1(y')$. One then shows that $z - z' \in \text{im } \pi_0$.

Moreover, we need to check that δ is a homomorphism (this is not completely obvious since the construction was not just a composition of homomorphisms).

Finally, one checks the exactness of the resulting long sequence (in two terms, $\ker \pi_2$ and $\text{coker } \pi_0$). We leave all these checks to the reader as an exercise. ■

Lemma 22.3.9 (3 × 3-lemma). Let the rows of the commutative diagram

$$\begin{array}{ccccccc}
 & 0 & 0 & 0 & & & \\
 & \downarrow & \downarrow & \downarrow & & & \\
 0 & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \longrightarrow 0 \\
 & \downarrow & \text{C} & \downarrow f & \text{C} & \downarrow & \\
 0 & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \longrightarrow 0 \\
 & \downarrow & \text{C} & \downarrow g & \text{C} & \downarrow & \\
 0 & \longrightarrow & \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \longrightarrow 0 \\
 & \downarrow & & \downarrow & & \downarrow & \\
 0 & & 0 & & 0 & &
 \end{array} \tag{22.3.15}$$

be exact. Then:

1. if the central and one of the side columns are exact, then the remaining column is exact too;
2. if the two side columns are exact and the middle one is a complex, i.e. $g \circ f = 0$, then the middle column is exact.

Proof. Exercise. ■

23 Cohomologies of Differential Forms

23.1 De Rham cohomology

Definition 23.1.1 (de Rham cohomology). For a smooth manifold M , consider the sequence of vector spaces of differential forms, called the *de Rham (cochain) complex*,

$$0 \rightarrow \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \Omega^2(M) \rightarrow \cdots \quad (23.1.1)$$

and define the spaces of closed forms, exact forms, and de Rham cohomology groups (in fact they are real vector spaces) respectively as

$$Z^p = \ker \left(d|_{\Omega^p(M)} \right), \quad B^p = \text{im} \left(d|_{\Omega^{p-1}(M)} \right), \quad H_{\text{dR}}^p(M) = Z^p / B^p. \quad (23.1.2)$$

This is possible because $d^2 = 0$, i.e. $B^p \subset Z^p$. Thus the sequence above is exact iff all de Rham cohomologies vanish.

Thus de Rham cohomology counts non-exact closed differential forms.

Proposition 23.1.2. 1. If M consists of l connected components, then $H_{\text{dR}}^0(M) = Z^0(M) = \mathbb{R}^l$.
2. If $\dim M = n$, then $H_{\text{dR}}^{>n}(M) = 0$.

Proof. Exercise. ■

Theorem 23.1.3 (Poincaré lemma). $H_{\text{dR}}^p(\mathbb{R}^n) = H_{\text{dR}}^p(\mathbb{R}^{n-1})$. More generally, for any manifold M ,

$$H_{\text{dR}}^\bullet(M \times \mathbb{R}) = H_{\text{dR}}^\bullet(M). \quad (23.1.3)$$

By induction, $H_{\text{dR}}^p(\mathbb{R}^n) = 0$ for $p > 0$.

Proof. We present a proof from [25] that uses the general ideas of homotopy operators that will be useful to us later. Let $\pi : M \times \mathbb{R} \rightarrow M$ be the projection on the first factor and $s : M \rightarrow M \times \mathbb{R}^1$ the zero section (or in fact any section). Then we have the corresponding pullback maps on differential forms:

$$\pi^* : \Omega^\bullet(M) \rightarrow \Omega^\bullet(M \times \mathbb{R}^1), \quad s^* : \Omega^\bullet(M \times \mathbb{R}^1) \rightarrow \Omega^\bullet(M). \quad (23.1.4)$$

Note that $\pi \circ s = 1$ and thus $s^* \circ \pi^* = 1$. Also, both s and π send closed forms to closed forms, which means that they induce well-defined maps on corresponding cohomology groups, which we will denote by the same symbols.

Let x denote the points of M and t the points of \mathbb{R}^1 . Every differential form on $M \times \mathbb{R}^1$ can be uniquely decomposed into a sum of the following types of forms:

$$f(x, t) \cdot (\pi^* \omega), \quad f(x, t) \cdot (\pi^* \omega) \wedge dt, \quad (23.1.5)$$

where ω is a form on M and $f(x, t)$ is a real-valued function on $M \times \mathbb{R}$ with $x \in M$.

Define the operator $K : \Omega^p(M \times \mathbb{R}^1) \rightarrow \Omega^{p-1}(M \times \mathbb{R}^1)$ by its action on the two kinds of forms from above,

$$f \cdot \pi^* \omega \mapsto 0, \quad f \cdot \pi^* \omega \wedge dt \mapsto (\pi^* \omega) \int_0^t f. \quad (23.1.6)$$

In other words, this operator integrates indefinitely over dt .

Let us now show that K is a *homotopy operator* between $\pi^* \circ s^*$ and the identity, which means that

$$\text{id} - \pi^* \circ s^* = \pm(dK \pm Kd), \quad (23.1.7)$$

where the precise arrangements of signs is irrelevant.

First let $\alpha = f(x, t) \cdot (\pi^* \omega)$ with $\deg \omega = p$ and compute

$$(1 - \pi^* s^*)\alpha = f(x, t)\pi^* \omega - f(x, 0)\pi^* \omega, \quad (23.1.8)$$

$$\begin{aligned} (dK - Kd)\alpha &= -Kd\alpha = -K(fd\pi^* \omega + (-1)^p df \wedge \pi^* \omega) = \\ &= (-1)^{p-1} \int_0^t \frac{\partial f}{\partial t} \pi^* \omega = (-1)^{p-1} (f(x, t) - f(x, 0)) \pi^* \omega. \end{aligned} \quad (23.1.9)$$

Therefore $(1 - \pi^* s^*)\alpha = (-1)^{p-1}(dK - Kd)\alpha$.

Now, for forms of the second type, $\alpha = f(x, t) \cdot (\pi^* \omega) \wedge dt$, we have

$$d\alpha = f\pi^* d\omega \wedge dt + (-1)^{p-1} \partial_x f \pi^* \omega \wedge dx \wedge dt, \quad (23.1.10)$$

$$s^* dt = 0 \rightarrow (1 - \pi^* s^*)\alpha = \alpha, \quad (23.1.11)$$

$$Kd\alpha = \left(\int_0^t f \right) \pi^* d\omega + (-1)^{p-1} \left(\int_0^t \partial_x f \right) \pi^* \omega \wedge dx, \quad (23.1.12)$$

$$dK\alpha = \left(\int_0^t f \right) \pi^* d\omega + (-1)^{p-1} \left(\int_0^t \partial_x f \right) \pi^* \omega \wedge dx + (-1)^{p-1} f \pi^* \omega \wedge dt, \quad (23.1.13)$$

so that $(dK - Kd)\alpha = (-1)^{p-1}\alpha$. In all cases we find

$$1 - \pi^* \circ s^* = (-1)^{p-1}(dK \pm Kd) \text{ on } \Omega^p(M \times \mathbb{R}). \quad (23.1.14)$$

It turns out that having a homotopy operator immediately allows us to relate cohomologies of different degrees to each other. Indeed, $dK \pm Kd$ maps closed forms to exact forms and therefore induces zero in cohomology (i.e. maps all cohomology equivalence classes to the trivial ones). In other words, the existence of such K implies that $\pi^* \circ s^* = 1$ in $H_{\text{dR}}^p(M)$. Therefore π^* and s^* are inverses of each other on cohomology:

$$H_{\text{dR}}^p(M \times \mathbb{R}) \cong H_{\text{dR}}^p(M). \quad (23.1.15)$$

■

Corollary 23.1.4. *De Rham cohomology is a homotopy invariant. In other words, if two smooth maps $f, g : M \rightarrow N$ are homotopic, then their actions in cohomology coincide:*

$$f \sim g \implies f_{\text{dR}}^* = g_{\text{dR}}^*. \quad (23.1.16)$$

Proof. We have the homotopy $H : M \times [0, 1] \rightarrow N$. Let $s_{0,1} : M \rightarrow M \times [0, 1]$ be two constant sections given by $s_i(m) = (m, i)$, $i = 0, 1$. Then $f = H \circ s_0$ and $g = H \circ s_1$. Their pullback actions are thus

$$f^* = s_0^* H^*, \quad g^* = s_1^* H^*. \quad (23.1.17)$$

However, we have shown in the proof of Poincaré lemma that $s_i^* = (\pi^*)^{-1}$ in de Rham cohomology (where $\pi : M \times [0, 1] \rightarrow M$ is the projection), regardless of the specific section. Thus

$$s_0^* = s_1^* \implies f^* = g^* \text{ in } H_{\text{dR}}^\bullet. \quad (23.1.18)$$

■

23.2 Mayer-Vietoris sequence

Suppose $M = U \cup V$ with U, V open. We have the inclusions

$$U \cap V \xrightarrow[i_1]{i_0} U \sqcup V \rightarrow M \quad (23.2.1)$$

Applying the contravariant functor Ω^\bullet , we get the sequence of restrictions of forms

$$\Omega^\bullet(M) \rightarrow \Omega^\bullet(U) \oplus \Omega^\bullet(V) \xrightarrow[i_1^*]{i_0^*} \Omega^\bullet(U \cap V). \quad (23.2.2)$$

By taking the difference of the last two maps, we obtain the *Mayer-Vietoris (MV) sequence*

$$0 \rightarrow \Omega^\bullet(M) \rightarrow \Omega^\bullet(U) \oplus \Omega^\bullet(V) \xrightarrow{\text{difference}} \Omega^\bullet(U \cap V) \rightarrow 0 \quad (23.2.3)$$

Proposition 23.2.1. *The MV sequence is commutative (i.e. the horizontal maps introduced above commute with applications of d) and exact.*

Proof. Commutativity is clear because d is local and thus commutes with restrictions. The only nontrivial part is the surjectivity of the difference map. Consider a POU $\{\chi_U, \chi_V\}$ subordinate to the open covering $\{U, V\}$ of M . Then any differential form $\omega \in \Omega^\bullet(U \cap V)$ can be decomposed as

$$\omega = \underbrace{\chi_U \omega}_{\Omega^\bullet(V)} - \underbrace{(-\chi_V)\omega}_{\Omega^\bullet(U)}, \quad (23.2.4)$$

which proves surjectivity. \blacksquare

Consider a general exact sequence of differential complexes

$$0 \rightarrow A^\bullet \rightarrow B^\bullet \rightarrow C^\bullet \rightarrow 0, \quad (23.2.5)$$

which is simply a shortened notation for the *commutative diagram*

$$\begin{array}{ccccccc} & & \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & A^{p+1} & \xrightarrow{f} & B^{p+1} & \xrightarrow{g} & C^{p+1} \longrightarrow 0 \\ & & d \uparrow & & d \uparrow & & d \uparrow \\ 0 & \longrightarrow & A^p & \xrightarrow{f} & B^p & \xrightarrow{g} & C^p \longrightarrow 0 \\ & & \uparrow & & \uparrow & & \uparrow \end{array} \quad (23.2.6)$$

in which every row is exact and $d^2 = 0$ in every column. The cohomology groups of each complex are again defined as $\ker d_p / \text{im } d_{p-1}$.

Essentially by the snake lemma 22.3.8, this induces a long exact sequence of cohomology groups:

$$\cdots \rightarrow H^p(A) \xrightarrow{f^*} H^p(B) \xrightarrow{g^*} H^p(C) \xrightarrow{\delta} H^{p+1}(A) \rightarrow H^{p+1}(B) \rightarrow H^{p+1}(C) \xrightarrow{\delta} H^{p+2}(A) \rightarrow \cdots \quad (23.2.7)$$

Namely, by surjectivity of g , for every closed $c \in C^p$, there exists a $b \in B^p$ such that $g(b) = c$. Using commutativity, $g(db) = d(gb) = dc = 0$, and by exactness, there exists an $a \in A^{p+1}$ such that $db = f(a)$. This a is easily checked to be closed $da = 0$. Then we define $\delta([c]) = [a] \in H^{p+1}(A)$. Some diagram chasing shows that this definition is independent of the choices made. We will discuss this sequence in detail later.

Applying this to the short **MV** sequence, we obtain the *long exact MV sequence*

$$\cdots \rightarrow H^p(M) \rightarrow H^p(U) \oplus H^p(M) \rightarrow H^p(U \cap V) \rightarrow H^{p+1}(M) \rightarrow H^{p+1}(U) \oplus H^{p+1}(V) \rightarrow H^{p+1}(U \cap V) \rightarrow \cdots \quad (23.2.8)$$

Retracing the construction of the connecting homomorphism, we find that

$$\delta^*([\omega]) = \begin{cases} [-d(\chi_V \omega)], & \text{on } U, \\ [d(\chi_U \omega)], & \text{on } V. \end{cases} \quad (23.2.9)$$

Example 23.2.2 (de Rham cohomology of \mathbb{S}^1). Let $M = \mathbb{S}^1$ and let U, V be a covering of the circle by two overlapping open arcs. Then we have the exact sequence

$$0 \rightarrow H^0(\mathbb{S}^1) \xrightarrow{\psi} H^0(U \sqcup V) \xrightarrow{\partial} H^0(U \cap V) \xrightarrow{\delta} H^1(\mathbb{S}^1) \rightarrow H^1(U \sqcup V) \rightarrow \cdots \quad (23.2.10)$$

in which by Proposition 23.1.2 and Poincaré lemma 23.1.3

$$H^0(\mathbb{S}^1) = \mathbb{R}, \quad H^0(U \sqcup V) = \mathbb{R}^2, \quad H^0(U \cap V) = \mathbb{R}^2, \quad H^1(U \sqcup V) = 0, \quad (23.2.11)$$

thus we get the exact sequence

$$0 \rightarrow \mathbb{R} \xrightarrow{\psi} \mathbb{R}^2 \xrightarrow{\partial} \mathbb{R}^2 \xrightarrow{\delta} H^1(\mathbb{S}^1) \rightarrow 0, \quad (23.2.12)$$

where

$$\partial(a, b) = (a - b, a - b). \quad (23.2.13)$$

Therefore by the rank-nullity theorem 22.1.7,

$$1 - 2 + 2 - \dim H^1(\mathbb{S}^1) = 0 \Rightarrow H^1(\mathbb{S}^1) = \mathbb{R}. \quad (23.2.14)$$

Alternatively, we can compute it as follows

$$\begin{aligned} \dim H^1(\mathbb{S}^1) &= \dim \text{im } \delta = \dim \mathbb{R}^2 - \dim \ker \delta = \\ &= 2 - \dim \text{im } \partial = 2 - (2 - \dim \ker \partial) = \dim \ker \partial = \dim \text{im } \psi = 1 \end{aligned} \quad (23.2.15)$$

since δ is surjective and ψ is injective.

Exercise 23.1. Using the **MV** sequence and homotopy invariance, show that $H_{\text{dR}}^k(\mathbb{S}^n) = H_{\text{dR}}^{k-1}(\mathbb{S}^{n-1})$ for $k \geq 2$, and by induction compute all de Rham cohomologies of any sphere

\mathbb{S}^n :

$$n > 0 : H_p(\mathbb{S}^n) = \begin{cases} \mathbb{Z}, & m = 0, n \\ 0, & \text{otherwise.} \end{cases} \quad (23.2.16)$$

Exercise 23.2. Using the homotopy invariance of de Rham cohomology, we can compute for example

$$H_{\text{dR}}^P(\mathbb{R}^n \setminus \{0\}) = H_{\text{dR}}^P(\mathbb{S}^{n-1}), \quad (23.2.17)$$

$$H_{\text{dR}}^1(\text{Möbius band}) = H_{\text{dR}}^1(\mathbb{S}^1) = \mathbb{Z}. \quad (23.2.18)$$

23.3 De Rham cohomology with compact support

Definition 23.3.1 (de Rham cohomology with compact support). Define $\Omega_c^\bullet(M)$ as the complex of differential forms with compact support on M . Then the cohomology groups $H_c^\bullet(M)$ of differential forms with compact support are defined just like before.

Proposition 23.3.2. Ω_c^\bullet is not a functor with respect to arbitrary smooth maps because pullbacks need not preserve compact supports. However:

1. Ω_c^\bullet is a contravariant functor under proper maps;
2. Ω_c^\bullet is a covariant functor under inclusions of open sets. Namely, if $i : U \hookrightarrow M$ is an inclusion of an open set, then $i_* : \Omega_c^\bullet(U) \hookrightarrow \Omega_c^\bullet(M)$ is the map that extends forms with compact support in U by zero to all of M .

Usually we will interpret Ω_c^\bullet as its covariant version.

A covering by two sets U, V with maps

$$U \cap V \xrightleftharpoons[i_1]{i_0} U \sqcup V \rightarrow M \quad (23.3.1)$$

as before gives rise to a sequence of forms with compact support:

$$\Omega_c^\bullet(U \cap V) \xrightarrow{j} \Omega_c^\bullet(U) \oplus \Omega_c^\bullet(V) \xrightarrow{\text{sum}} M, \quad (23.3.2)$$

where j is the “signed inclusion” $\omega \mapsto (-i_*\omega, i_*\omega)$.

Proposition 23.3.3. The MV sequence of forms with compact support

$$0 \rightarrow \Omega_c^\bullet(U \cap V) \rightarrow \Omega_c^\bullet(U) \oplus \Omega_c^\bullet(V) \rightarrow \Omega_c^\bullet(M) \rightarrow 0 \quad (23.3.3)$$

is exact.

Proof. The least trivial step is the last one. Let $\omega \in \Omega_c^\bullet(M)$. Then ω is the image of $(\chi_U \omega, \chi_V \omega)$. The form $\chi_U \omega$ indeed has compact support in U , therefore the map $\Omega_c^\bullet(U) \oplus \Omega_c^\bullet(V) \rightarrow \Omega_c^\bullet(M)$ is surjective. Note that unlike in the previous MV sequence, here we multiply by χ_U to get a form on U . ■

This sequence, just like the last one, produces a long exact sequence in cohomology:

$$\cdots \rightarrow H_c^p(U \cap V) \rightarrow H_c^p(U) \oplus H_c^p(V) \rightarrow H_c^p(M) \rightarrow H_c^{p+1}(U \cap V) \rightarrow H_c^{p+1}(U) \oplus H_c^{p+1}(V) \rightarrow H_c^{p+1}(M) \rightarrow \cdots \quad (23.3.4)$$

Obviously, for compact manifolds $H_c^\bullet = H_{\text{dR}}^\bullet$, so in this case we have two MV sequences in both directions!

Exercise 23.3. Use this exact sequence to show again that $H_c^0(\mathbb{S}^1) = H_c^1(\mathbb{S}^1) = \mathbb{R}$.

Theorem 23.3.4 (Poincaré lemma for compact supports). *For any manifold M ,*

$$H_c^\bullet(M \times \mathbb{R}) = H_c^{\bullet-1}(M). \quad (23.3.5)$$

In particular $H_c^p(\mathbb{R}^n) = 0$ for $p \neq n$ and $H_c^n(\mathbb{R}^n) = \mathbb{R}$.

Proof. We will use the same ideas and notation as in the first Poincaré lemma 23.1.3. Consider again the projection $\pi : M \times \mathbb{R} \rightarrow M$. The pullback π^* does not map nonzero forms into forms with compact support on $M \times \mathbb{R}$, however there is a push-forward map $\pi_* : \Omega_c^\bullet(M \times \mathbb{R}) \rightarrow \Omega_c^{\bullet-1}(M)$ called *integration along the fiber*. We define it separately on the two kinds of forms:

$$\pi_* : \quad f(x, t)\pi^*\omega \mapsto 0, \quad f(x, t)\pi^*\omega \wedge dt \mapsto \omega \cdot \int_{\mathbb{R}} f(x, t)dt. \quad (23.3.6)$$

One can check that $d\pi_* = \pi_* d$, therefore we have an induced map in cohomology $\pi_* : H_c^\bullet \rightarrow H_c^{\bullet-1}$.

To define a map in the reverse direction, choose $e = e(t)dt$ to be any form on \mathbb{R} of total integral 1 and define

$$e_* : \Omega_c^\bullet(M) \rightarrow \Omega_c^{\bullet+1}(M \times \mathbb{R}), \quad \omega \mapsto (\pi^*\omega) \wedge e. \quad (23.3.7)$$

This map clearly commutes with d and thus also descends to cohomology. From the definition is also follows that

$$\pi_* \circ e_* = 1 \text{ on } \Omega_c^\bullet(M). \quad (23.3.8)$$

To prove that $e_* \circ \pi_* = 1$ in cohomology as well, we introduce the new homotopy operator $K : \Omega_c^\bullet(M \times \mathbb{R}) \rightarrow \Omega_c^{\bullet-1}(M \times \mathbb{R})$ by

$$K : \quad f \cdot \omega \mapsto 0, \quad f \cdot \omega \wedge dt \mapsto \omega \cdot \int_{\mathbb{R}} f - \omega \cdot E(t) \int_{\mathbb{R}} f, \quad \text{where } E(t) = \int_{-\infty}^t e \quad (23.3.9)$$

Now it remains to show that

$$1 - e_* \pi_* = (-1)^{p-1}(\mathrm{d}K - K\mathrm{d}) \text{ on } \Omega_c^p(M \times \mathbb{R}). \quad (23.3.10)$$

For forms of the first kind, $\alpha = f \cdot \omega$, we have

$$(1 - e_* \pi_*)f \cdot \omega = f \cdot \omega \quad (23.3.11)$$

and

$$\begin{aligned} (\mathrm{d}K - K\mathrm{d})\alpha &= -K(f \mathrm{d}\omega + (-1)^p \partial_x f \omega \wedge dx + (-1)^p \partial_t f \omega \wedge dt) = \\ &= (-1)^{p-1} \left(\omega \cdot \int_{-\infty}^t \partial_t f - \underbrace{\omega \cdot E(t) \int_{\mathbb{R}} \partial_t f}_0 \right) = (-1)^{p-1} \alpha. \end{aligned} \quad (23.3.12)$$

For forms of the second type, $\alpha = f \pi^* \omega \wedge dt$, we have

$$(1 - e_* \pi_*)\alpha = f \omega \wedge dt - \omega \wedge e \cdot \int_{\mathbb{R}} f, \quad (23.3.13)$$

$$\begin{aligned} \mathrm{d}K(f \omega \wedge dt) &= \mathrm{d}\omega \cdot \int_{-\infty}^t f + (-1)^{p-1} \int_{-\infty}^t \partial_x f \omega \wedge dx + \\ &+ (-1)^{p-1} f \omega \wedge dt - E(t) \int_{\mathbb{R}} f \cdot \mathrm{d}\omega - (-1)^{p-1} \omega \left[e \int_{\mathbb{R}} f + dx \cdot E(t) \int_{\mathbb{R}} \partial_x f \right], \end{aligned} \quad (23.3.14)$$

$$\begin{aligned} K\mathrm{d}(f \omega \wedge dt) &= K \left(f \mathrm{d}\omega \wedge dt + (-1)^{p-1} \partial_x f \omega \wedge dx \wedge dt \right) = \\ &= \int_{-\infty}^t f \mathrm{d}\omega - \mathrm{d}\omega E(t) \int_{\mathbb{R}} f + (-1)^{p-1} \omega \wedge dx \left[\int_{-\infty}^t \partial_x f - \int_{\mathbb{R}} \partial_x f \right]. \end{aligned} \quad (23.3.15)$$

Therefore

$$(\mathrm{d}K - K\mathrm{d})f\omega \wedge dt = (-1)^{p-1} \left[f\omega \wedge dt - \omega \wedge e \cdot \int_{\mathbb{R}} f \right], \quad (23.3.16)$$

which proves the needed identity.

The existence of a homotopy operator proves that the maps π_* and e_* provide an isomorphism $H_c^\bullet(M \times \mathbb{R}) \cong H_c^{\bullet-1}(M)$.

$H_c^n(\mathbb{R}^n) \cong \mathbb{R}$ by iteration since $H_c^0(\text{point}) = H_{\text{dR}}^0(\text{point}) = \mathbb{R}$. ■

Remark 23.3.5. This theorem shows that cohomology with compact supports is *not* a homotopy invariant (the fiber \mathbb{R} in $M \times \mathbb{R}$ can be retracted by a homotopy), unlike the usual de Rham cohomology.

23.4 Finite dimensionality of de Rham cohomology

We have already seen that the MV sequence combined with the Poincaré lemma is a powerful tool that is sufficient to compute cohomologies of nontrivial manifolds like spheres. Now we will show that these two tools are in fact sufficient for arbitrary manifolds, as long as we choose the right open covering.

Definition 23.4.1 (Good covers). An open covering $\mathcal{U} = \{U_\alpha\}_\alpha$ of an n -dimensional smooth manifold M is called a good cover if all nonempty finite intersections $U_{\alpha_1 \dots \alpha_k} = U_{\alpha_1} \cap \dots \cap U_{\alpha_k}$ are diffeomorphic to \mathbb{R}^n (i.e. are contractible).

Proposition 23.4.2. *Every smooth manifold has a good cover.*

Proof. We have shown earlier that every manifold admits a Riemannian metric, and every Riemannian manifold can be covered by geodesically convex neighborhoods whose intersections are also geodesically convex (Theorem 18.2.15). Therefore such a cover is a good cover. ■

Recall that an open covering \mathcal{V} is a *refinement* of \mathcal{U} if every V_β is contained in some U_α , and we write $\mathcal{U} \leq \mathcal{V}$.

Corollary 23.4.3. *Every open covering of a manifold has a good refinement.*

Definition 23.4.4 (Directed and cofinal sets). A directed set is a set I with a relation \leq satisfying:

1. (reflexivity) $a \leq a$ for all a ;
2. (transitivity) $a \leq b, b \leq c$ implies $a \leq c$;
3. (upper bound) $\forall a, b \in I : \exists c \in I : a \leq c$ and $b \leq c$.

A subset $J \subset I$ is *cofinal* in I if $\forall i \in I \exists j \in J : i \leq j$. Such a J is automatically a directed set too.

Note that $\neg(a \leq b)$ here does not imply $b \leq a$. So, even though the relation is defined for all pairs (a, b) , it is not true that at least one of $a \leq b$ or $b \leq a$ always holds.

The set of all open coverings of M is a directed set with the refinement relation, since any two covers have a common refinement. With this, we can restate the last Corollary as follows.

Corollary 23.4.5. *The set of all good covers is cofinal in the set of all open coverings of M .*

Proposition 23.4.6. *If a smooth manifold M has a finite good cover, then its de Rham cohomology (with compact supports or not) is finite-dimensional. In particular this holds for all compact manifolds.*

Proof. From the **MV** sequence

$$\cdots \rightarrow H_{\text{dR}}^{q-1}(U \cap V) \xrightarrow{\delta} H^q(U \cap V) \xrightarrow{r} H^q(U) \oplus H^q(V) \rightarrow \cdots \quad (23.4.1)$$

we have by the rank-nullity theorem

$$H_{\text{dR}}^q(U \cup V) \cong \ker r \oplus \text{im } r \cong \text{im } \delta \oplus \text{im } r. \quad (23.4.2)$$

Thus if $H^q(U)$, $H^q(V)$ and $H^{q-1}(U \cap V)$ are finite-dimensional, then so is $H^q(U \cup V)$. We proceed by induction in the size of the good cover. The base of induction holds by the Poincaré lemma. Suppose that the cohomology of any manifold having a good cover with at most p elements is finite-dimensional. Consider a manifold with a good cover $\{U_0, \dots, U_p\}$. Now $(U_0 \cup \dots \cup U_{p-1}) \cap U_p$ has a good cover with p elements, namely $\{U_{0,p}, U_{1,p}, \dots, U_{p-1,p}\}$. Thus by hypothesis, the q th cohomology of $(U_0 \cup \dots \cup U_{p-1})$, U_p and $(U_0 \cup \dots \cup U_{p-1}) \cap U_p$ are finite-dimensional. Using the above result for a union of two open sets, we conclude that so is the cohomology of $U_0 \cup \dots \cup U_p$. ■

23.5 Hodge theory

23.6 Cohomology of Lie groups

23.7 Dolbeault cohomology

Similar to de Rham cohomology, we can define cohomology of complex differential forms.

Definition 23.7.1 (Dolbeault cohomology). For M a complex manifold, define the Dolbeault

cohomology consisting of $\bar{\partial}$ -closed forms:

$$H^{p,q}(M) = \frac{\ker \bar{\partial}|_{\Omega^{p,q}(M)}}{\text{im } \bar{\partial}|_{\Omega^{p,q-1}(M)}}. \quad (23.7.1)$$

There is also an analog of the Poincaré lemma for $\bar{\partial}$, but at first only over one-dimensional complex manifolds.

Lemma 23.7.2 ($\bar{\partial}$ -Poincaré lemma in one variable). *Let $D = \{z \in \mathbb{C} \mid |z| < \epsilon\}$ and $U \subset \mathbb{C}$ such that $\overline{D} \subset U$. Let $\omega \in \Omega^{0,1}(U)$. Then ω is $\bar{\partial}$ -exact:*

$$\omega = \bar{\partial}g, \quad g(z) = \frac{1}{2\pi i} \int_D \frac{d\xi \wedge \omega(\xi)}{\xi - z}. \quad (23.7.2)$$

Proof. Let $\omega = f d\bar{z}$ with $f \in C^\infty(U, \mathbb{C})$ and $z_0 \in D$. Let $\chi : D \rightarrow \mathbb{R}$ be a bump with compact support around z_0 and inside D : $\chi|_V = 1$ for $z_0 \in V \subset D$. Set

$$f_1 = \chi \cdot f, \quad f_2 = f - \chi \cdot f, \quad (23.7.3)$$

so that $f = f_1 + f_2$. Consider the integrals

$$g_i(z) = \frac{1}{2\pi i} \int_D \frac{f_i(\xi) d\xi \wedge d\bar{\xi}}{\xi - z}, \quad i = 1, 2. \quad (23.7.4)$$

Since $f_2|_V = 0$, g_2 is obviously well defined for $z \in V$. g_1 is also well defined because the $1/(z - \xi)$ singularity is summable in two dimensions. Now we compute $\bar{\partial}g$. Since $1/(z - \xi)$ is holomorphic on the support of f_2 , we have

$$\bar{\partial}g_2 = 0. \quad (23.7.5)$$

On the other hand, f_1 has compact support and thus we can treat $1/(\xi - z)$ as a distribution acting on the test function f_1 . From the general $\bar{\partial}$ -formula (which is nothing but an application of the Stokes theorem)

$$f(z) = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(\xi) d\xi}{\xi - z} + \frac{1}{2\pi i} \int_D \frac{\bar{\partial}f(\xi) d\xi \wedge d\bar{\xi}}{\xi - z}, \quad (23.7.6)$$

we know the distributional identity

$$\bar{\partial}_z \frac{1}{z - \xi} = \pi \delta(z - \xi). \quad (23.7.7)$$

Thus

$$\bar{\partial}g_1 = -\frac{1}{2i} \int_D f_1(\xi) \delta(\xi - z) \cdot (-2i) dV(\xi) = f_1(z) = f(z). \quad (23.7.8)$$

$g = g_1 + g_2$ is independent of χ and satisfies the theorem. ■

Theorem 23.7.3 ($\bar{\partial}$ -Poincaré lemma in several variables (Dolbeault-Grothendieck)). *Let D be a polydisc $D \subset \mathbb{C}^n$ (a polydisc is a direct product of any n open discs in \mathbb{C} , including \mathbb{C} itself). If $\omega \in \Omega^{p,q}(D)$ is $\bar{\partial}$ -closed and $q > 0$, then it is $\bar{\partial}$ -exact:*

$$\omega = \bar{\partial}\eta, \quad \eta \in \Omega^{p,q-1}(D). \quad (23.7.9)$$

In other words, $H^{p,q}(D) = 0$ for $q > 0$.

Proof. For bounded polydiscs this can be reduced to an inductive application of the one-dimensional lemma. In the unbounded case the proof requires some more careful analysis, see e.g. [24, Prop. 1.3.8 and Cor. 1.3.9]. ■

The following proposition is obvious based on its de Rham analogue.

Proposition 23.7.4. *If $\dim_{\mathbb{C}} M = n$ and $p + q > 2n$, then $H^{p,q}(M) = 0$.*

Notably, it is not true that $H^{0,0}(M)$ simply counts connected components of M , since there may be non-constant holomorphic functions on M .

Computing Dolbeault cohomology in general requires more sophisticated tools than de Rham cohomology, since there is no analog of the MV sequence here (we can't multiply by bump functions and still have $\bar{\partial}$ -closed forms). The most efficient way of studying these groups is through sheaf cohomology.

Example 23.7.5 (Dolbeault cohomology of the Riemann sphere). Here we compute the Dolbeault cohomology of the Riemann sphere $\mathbb{C}P^1 = \mathbb{S}^2$.

1. $H^{0,0}(\mathbb{C}P^1) = \mathbb{C}$ (the only holomorphic functions on the entire sphere are the constant functions, by the Liouville theorem);
2. $H^{0,1}(\mathbb{C}P^1)$ consists of 1-forms $f d\bar{z}$ (all of which are $\bar{\partial}$ -closed for any $f \in C^\infty(\mathbb{C}P^1, \mathbb{C})$) that are not exact, i.e. $f \neq \partial_{\bar{z}}g$ for any $g \in C^\infty(\mathbb{C}P^1, \mathbb{C})$. Every such f can be identified with a function on \mathbb{C} such that the limit of $f\bar{z}^2$ as $z \rightarrow \infty$ exists. By Poincaré lemma there exists a g such that $f = \partial_{\bar{z}}g$ and clearly it follows that $\lim_{z \rightarrow \infty} (\bar{z}g)$ exists, which makes g a well-defined function on the sphere. Therefore $H^{0,1}(\mathbb{C}P^1) = 0$;
3. $H^{1,0}(\mathbb{C}P^1)$ consists of *holomorphic* 1-forms $f dz$, which means that f has to be holomorphic such that $\lim_{z \rightarrow \infty} z^2 f$ exists. However, the only holomorphic function with such a limit is zero. Therefore $H^{1,0}(\mathbb{C}P^1) = 0$;

4. $H^{1,1}(\mathbb{C}P^1)$ consists of 2-forms $f dz \wedge d\bar{z}$. Such a form is exact iff $\int_{\mathbb{C}P^1} f dz \wedge d\bar{z} = 0$.

This is exactly one linear constraint, which means that $H^{1,1}(\mathbb{C}P^1) = \mathbb{C}$.

With more tools, one can show that $H^{p,q}(\mathbb{C}P^n)$ is in general zero for $p \neq q$ and $H^{p,p}(\mathbb{C}P^n) = \mathbb{C}$ for $p = 0, \dots, n$.

24 Homology of Topological Spaces

24.1 Simplicial homology

Simplicial homology is the simplest homology theory in that it can be formulated entirely within basic set theory, despite the fact that in the applications interesting to us the elements these sets will usually have a topological or geometric interpretation. More details on simplicial complexes can be found in [26].

Definition 24.1.1 (Simplicial complexes). Let E be a set. A finite subset $s = \{p_0, \dots, p_n\}$ of E is called an n -simplex (or an n -dimensional simplex). Subsets of s are called its *faces*. A *simplicial complex* is a pair $K = (E, S)$ where S is a collection of simplices in E such that:

1. $\forall e \in E, \{e\} \in S$ (every 0-simplex is in the complex);
2. If $s \in S$ and $s' \subset s$ is non-empty, then $s' \in S$ (all faces of a simplex are also in the complex).

0-simplices are called *vertices* of K , 1-simplices are called *edges*. The complex is called n -dimensional if it contains an n -simplex but no $(n+1)$ -simplices. A *subcomplex* $L \subset K$ is a subset of its simplices which contains with each simplex also its faces.

A 1-dimensional complex is called a graph. $K = (E, S)$ is called *finite* if E is finite and *locally finite* if each vertex is contained in finitely many simplices.

The sub-complex $K^n = (E, S^n)$ of K with $S^n = \{s \in S \mid \dim s \leq n\}$ is called the *n-skeleton* of K .

Example 24.1.2. 1. Let $\mathcal{U} = \{U_i\}_{i \in J}$ be a covering of a set X by non-empty subsets. For a finite subset $E \subset J$ let $U_E = \bigcap_{j \in E} U_j$ and let $E(J) = \{E \subset J \mid U_E \neq \emptyset\}$. Then $(J, E(J))$ is a simplicial complex called the *nerve* $N(\mathcal{U})$ of the covering \mathcal{U} .
2. Let (P, \leq) be a poset. The simplicial complex (P, S_P) associated to a poset has as simplices the totally ordered finite subsets of P .
3. If $K = (E, S)$ is a simplicial complex, define the inclusion partial order on S . The simplicial complex K' associated to the poset (S, \subset) is called the *barycentric subdivision* of K . We will visualize this in the next example.

Definition 24.1.3 (Barycentric coordinates). Let $K = (E, S)$ be a simplicial complex. Denote by $|K|$ the set of functions $\alpha : E \rightarrow [0, 1]$ such that:

1. $\{e \in E \mid \alpha(e) > 0\}$ is a simplex of K ;
2. $\sum_{e \in E} \alpha(e) = 1$.

Therefore $|K| \subset [0, 1]^E$ and induces a topology from the product topology. We denote this topological space by $|K|_p$. The values $\{\alpha(e)\}_{e \in E}$ are called *barycentric coordinates* of α . For $s \in S$, the set $\Delta(s) = \{\alpha \in |K| \mid \forall e \notin s, \alpha(e) = 0\} \subset [0, 1]^E$ is called a *standard simplex*. Then $|K|$ is the union of all $\Delta(s)$ and we write $|K|_c$ for $|K|$ with the quotient

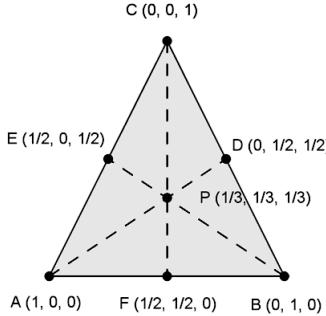


Fig. 24.1: Barycentric coordinates of a 2-simplex. Here, the simplex is $s = \langle A, B, C \rangle$ and the triples of numbers next to each point $\alpha \in |s|$ in this triangle correspond to the list of values $(\alpha(A), \alpha(B), \alpha(C))$. The values of α thus represent the proportions in which A, B, C have to be “mixed” to get a point inside this geometric realization of the simplex.

topology induced from $([0, 1]^E)^S$ by the canonical map $\bigsqcup_{s \in S} \Delta(s) \rightarrow |K|$.

Example 24.1.4. The space of barycentric coordinates for a 2-simplex $\langle A, B, C \rangle$ can be represented by the triangle in Fig. 24.1. Its barycentric subdivision is a complex whose simplices are the inclusion-ordered sets of subsets of $\{A, B, C\}$. For instance, the simplex $\langle P \rangle = \{\{A, B, C\}\}$ is a 0-simplex represented by the center of the triangle; the simplex $\langle A, P \rangle = \{\{A\}, \{A, B, C\}\}$ is the 1-simplex connecting A to the center, and $\langle A, F \rangle = \{\{A\}, \{A, B\}\}$ is the 1-simplex connecting A to F ; the 2-simplex $\langle A, F, P \rangle = \{\{A\}, \{A, B\}, \{A, B, C\}\}$ is the triangle whose vertices are A, F , and the center P . There are six 2-simplices in this subdivision, for instance $\langle A, E, P \rangle = \{\{A\}, \{A, C\}, \{A, B, C\}\}$.

Proposition 24.1.5. *The identity map $|K|_c \rightarrow |K|_p$ is a homotopy equivalence.*

Proof. See [26, Prop. 8.1.4 and 13.2.2]. ■

In the future we write $|K| = |K|_c$ and call this space the *geometric realization* of K . For a simplex s we define the *closed simplex* $|s| \subset |K|$ as $|s| = \{\alpha \in |K| \mid \alpha(e) \neq 0 \Rightarrow e \in s\}$ and the *open simplex* as $\langle s \rangle = \{\alpha \in |K| \mid \alpha(e) \neq 0 \Leftrightarrow e \in s\}$. The *combinatorial boundary* of $|s|$ is $\partial|s| = |s| \setminus \langle s \rangle$.

Definition 24.1.6 (Standard simplices in \mathbb{R}^k). Let x_0, \dots, x_n be affinely independent (i.e.

$\sum_i \alpha_i x_i = 0$ and $\sum_i \alpha_i = 0$ imply $\alpha_j = 0$). Then the simplex spanned by them is

$$\langle x_0, \dots, x_n \rangle = \left\{ \sum_i \alpha_i x_i \mid \alpha_i \geq 0, \sum_i \alpha_i = 1 \right\}. \quad (24.1.1)$$

Definition 24.1.7. If $K = (E, S)$ is a simplicial complex and $\{x_e \mid e \in E\}$ a set of points in \mathbb{R}^k , and if a map

$$f : |K| \rightarrow \mathbb{R}^k, \quad \alpha \mapsto \sum_{e \in E} \alpha(e) x_e \quad (24.1.2)$$

is an embedding (say, topological), then $f(|K|)$ is called a simplicial polyhedron in \mathbb{R}^k of type K , or a polyhedral realization of K in \mathbb{R}^k .

Definition 24.1.8 (Triangulation). Let X be a topological space. A triangulation of X is a triple (X, K, f) where K is a simplicial complex K such that there is a homeomorphism $f : X \rightarrow |K|$.

It is known that all differentiable manifolds can be triangulated, and the triangulation can be chosen so that on each simplex it is a smooth embedding.

Definition 24.1.9 (Simplicial homology). For a simplicial complex $K = (E, S)$ define the group $C_p^\Delta(K)$ of *simplicial p-chains* as the free abelian group generated by the set of all p -simplices in K . That is, elements of $C_p^\Delta(K)$ are formal sums $\sum_s c_s s$, where s runs over all p -simplices in K , and only a finite number of coefficients $c_s \in \mathbb{Z}$ are nonzero at once.

Define the *boundary operator*

$$\partial : C_p^\Delta(K) \rightarrow C_{p-1}^\Delta(K), \quad \partial \langle e_0, \dots, e_p \rangle = \sum_{i=0}^p (-1)^i \langle e_0, \dots, \hat{e}_i, \dots, e_p \rangle, \quad (24.1.3)$$

where the hat means skipping a vertex. It is easy to check that $\partial^2 = 0$, therefore we have the *chain complex*

$$\cdots \rightarrow C_2^\Delta(K) \xrightarrow{\partial} C_1^\Delta(K) \xrightarrow{\partial} C_0^\Delta(K) \rightarrow 0. \quad (24.1.4)$$

The group of p -cycles is $Z_p = \ker \partial|_{C_p^\Delta(M)}$ and the group of p -boundaries is $B_p = \text{im } \partial|_{C_{p+1}^\Delta(M)}$. The *simplicial homology groups* of K are defined as

$$H_p^\Delta(K) = Z_p / B_p. \quad (24.1.5)$$

Intuitively, H_p^Δ detects the “ p -dimensional holes” in the complex, because it consists of cycles that are not boundaries of anything $p + 1$ -dimensional.

Definition 24.1.10 (Orientations on simplices and complexes). An orientation on an n -simplex s is an ordering of its elements into an $(n + 1)$ -tuple written as

$$s = \langle e_0, \dots, e_n \rangle. \quad (24.1.6)$$

Orientations are considered to be equivalent if they differ by an even permutation of the vertices, i.e. we identify positively reordered tuples with each other.

In simplicial p -chains, sometimes the chain $-s$ is identified with s oriented the opposite way. An orientation on a complex $K = (E, S)$ is a partial order on E that induces an orientation on each simplex of K (i.e. orientations of faces of a simplex can be induced from the orientation of the simplex). Orientations are considered to be equivalent if they induce equivalent orientations on all simplices.

Example 24.1.11 (Simplicial homology of \mathbb{S}^1). A circle \mathbb{S}^1 can be triangulated by a regular n -gon with vertices e_0, \dots, e_{n-1} and oriented simplices $s_i = \langle e_i, e_{i+1} \rangle$, $0 \leq i \leq n - 1$ with the identification $e_n = e_0$. Clearly

$$Z_0 = C_0^\Delta = \mathbb{Z}^n, \quad C_1^\Delta = \left\{ \sum_i c_i s_i \mid c_i \in \mathbb{Z} \right\} \cong \mathbb{Z}^n, \quad B_1 = 0. \quad (24.1.7)$$

The chain complex is described by

$$\partial \langle e_i, e_{i+1} \rangle = \langle e_i \rangle - \langle e_{i+1} \rangle. \quad (24.1.8)$$

B_0 consists of 0-chains $\sum c_i \langle e_i \rangle$ such that $\sum c_i = 0$, thus $B_0 \cong \mathbb{Z}^{n-1}$. Finally, Z_1 consists of 1-chains $\sum_i c_i s_i$ such that $c_i - c_{i+1} = 0$ for all i , which means that $Z_1 \cong \mathbb{Z}$ generated by the cycle $\sum_i s_i$. In the end, we have

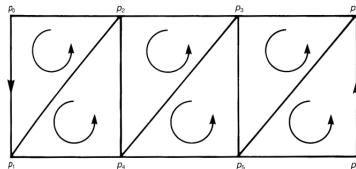
$$H_0^\Delta \cong \mathbb{Z}^n / \mathbb{Z}^{n-1} \cong \mathbb{Z}, \quad H_1^\Delta = Z_1 \cong \mathbb{Z}. \quad (24.1.9)$$

We notice that the ranks of these homology groups coincide with the dimensions of the corresponding de Rham cohomologies, see Example 23.2.2. We will later prove this as a general fact for all manifolds: the de Rham cohomology is isomorphic to the simplicial homology with real coefficients.

Exercise 24.1. If K is the tetrahedron (with 2-dimensional faces included), i.e. a triangulation of \mathbb{S}^2 , compute its homology groups:

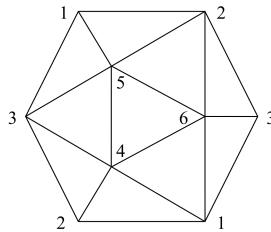
$$H_0^\Delta = \mathbb{Z}, \quad H_1^\Delta = 0, \quad H_2^\Delta = \mathbb{Z}. \quad (24.1.10)$$

Exercise 24.2. Triangulate the Möbius band as in the figure.



Show that $H_2^\Delta(K) = 0$ and $H_1^\Delta(K) = \mathbb{Z}$. In fact, the generator of H_1^Δ is the loop going from the bottom left to the top right corner in the diagram above.

Exercise 24.3. Triangulate the real projective plane $\mathbb{R}P^2$ (which is homeomorphic to a disk whose antipodal boundary points have been identified) as in the figure.



Show that $H_2^\Delta(K) = 0$ and $H_1^\Delta(K) = \mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$.

Exercise 24.4. Triangulate the Klein bottle and show that $H_1^\Delta(K) = \mathbb{Z} \oplus \mathbb{Z}_2$.

24.2 Homological algebra II: complexes

Definition 24.2.1 ((Co)chain complexes). A chain complex (C, d) in an abelian category C is a sequence of morphisms

$$\cdots \rightarrow C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \rightarrow \cdots \quad (24.2.1)$$

indexed by $n \in \mathbb{Z}$, such that $d_n \circ d_{n+1} = 0$.

Similarly, a cochain complex is a sequence

$$\cdots \rightarrow C^{n-1} \xrightarrow{d^{n-1}} C^n \xrightarrow{d^n} C^{n+1} \rightarrow \cdots \quad (24.2.2)$$

such that $d^{n+1} \circ d^n = 0$.

Note that there is no actual difference between chain and cochain complexes, since simply changing the enumeration $n \mapsto -n$ turns one into the other. Therefore all general results need to be proven only for chain complexes.

Definition 24.2.2 ((Co)cycles, (co)boundaries, (co)homologies of complexes). For a chain complex (\mathbf{C}, d) , n -cycles, n -boundaries, and homologies are defined as

$$Z_n(\mathbf{C}, d) = \ker d_n, \quad B_n(\mathbf{C}, d) = \text{im } d_{n+1}, \quad H_n(\mathbf{C}, d) = Z_n(\mathbf{C}, d)/B_n(\mathbf{C}, d). \quad (24.2.3)$$

Similarly, for a cochain complex (C^\bullet, d) , we have cocycles, coboundaries, and cohomologies:

$$Z^n(C^\bullet, d) = \ker d_n, \quad B^n(C^\bullet, d) = \text{im } d^{n-1}, \quad H^n(C^\bullet, d) = Z^n(C^\bullet, d)/B^n(C^\bullet, d). \quad (24.2.4)$$

Definition 24.2.3 (Chain map). If (\mathbf{B}, δ) and (\mathbf{C}, d) are two chain complexes, then a chain map $f : \mathbf{B} \rightarrow \mathbf{C}$ is a sequence of maps $f_n : B_n \rightarrow C_n$ such that $d_n \circ f_n = f_{n-1} \circ \delta_n$, i.e. such that the full diagram commutes:

$$\begin{array}{ccccccc} \cdots & \longrightarrow & B_{n+1} & \longrightarrow & B_n & \longrightarrow & B_{n-1} \longrightarrow \cdots \\ & & \downarrow f_{n+1} & \curvearrowright & \downarrow f_n & \curvearrowright & \downarrow f_{n-1} \\ \cdots & \longrightarrow & C_{n+1} & \longrightarrow & C_n & \longrightarrow & C_{n-1} \longrightarrow \cdots \end{array} \quad (24.2.5)$$

Similarly one defines cochain maps between cochain complexes.

Proposition 24.2.4. All chain complexes in an abelian category \mathcal{C} with chain maps between them comprise an abelian category called $\text{Comp}(\mathcal{C})$.

Proof. By the Freyd-Mitchell theorem, we only need to prove this for $\mathcal{C} = R\text{-Mod}$. For such a category it is obvious what the structure of an abelian category on complexes is: the zero object is the zero complex; direct sums are defined by summing in each degree; the kernel and cokernel of f are the complexes consisting of $\ker f_n$ and $\text{coker } f_n$; images and coimages coincide because they do so in \mathcal{C} . ■

Definition 24.2.5 (Positive complexes). A chain complex \mathbf{C} is called positive if $C_n = 0, n < 0$. All positive complexes form the full subcategory $\text{Comp}_{>0}(\mathcal{C})$ of $\text{Comp}(\mathcal{C})$:

$$\cdots \rightarrow C_n \rightarrow C_{n-1} \rightarrow \cdots \rightarrow C_1 \rightarrow C_0 \rightarrow 0. \quad (24.2.6)$$

A negative chain complex

$$0 \rightarrow C_0 \rightarrow C_{-1} \rightarrow \cdots \rightarrow C_{-n} \rightarrow C_{-n-1} \rightarrow \cdots \quad (24.2.7)$$

is identified with a positive cochain complex

$$0 \rightarrow C^0 \rightarrow C^1 \rightarrow \cdots \rightarrow C^n \rightarrow C^{n+1} \rightarrow \cdots \quad (24.2.8)$$

by setting $C^n = C_{-n}$ and $d^n = d_{-n}$.

Proposition 24.2.6. *If C is an abelian category, then the n -th homology $H_n : \text{Comp}(C) \rightarrow C$ is an additive (covariant) functor for all $n \in \mathbb{Z}$.*

Proof. Given a chain map $f : \mathbf{A} \rightarrow \mathbf{B}$, we need to construct a morphism $H_n(f) : H_n(\mathbf{A}) \rightarrow H_n(\mathbf{B})$. Since $H_n(\mathbf{A}) = Z_n(\mathbf{A})/B_n(\mathbf{A})$, we can try to take a z such that

$$f_n : Z_n(\mathbf{A}) \rightarrow B_n, \quad z \mapsto f_n(z) \quad (24.2.9)$$

and define

$$H_n(f) : H_n(\mathbf{A}) \rightarrow H_n(\mathbf{B}), \quad [z] \mapsto [f_n(z)], \quad (24.2.10)$$

where equivalence classes are taken w.r.t. to the quotients by $B_n(\mathbf{A})$ and $B_n(\mathbf{B})$ respectively. First we need to check that $f_n(z) \in Z_n(\mathbf{B})$. We will use the same symbol d_n for morphisms in both complexes. By the definition of a chain map, $d_n(f_n(z)) = f_{n-1}(d_n(z)) = 0$ since $z \in B_n(\mathbf{A})$.

Next we need to check correctness, i.e. independence of $H_n(f)([z])$ of the choice of representative. Let $z, z' \in Z_n(\mathbf{A}) : [z] = [z'] \Leftrightarrow z - z' \in B_n(\mathbf{A})$, then

$$\exists y \in A_{n+1} : d_{n+1}(y) = z - z' \implies d_{n+1}(f_{n+1}(y)) = f_n(z) - f_n(z'), \quad (24.2.11)$$

which means that $f_n(z) - f_n(z')$ is a boundary, proving what we wanted.

To check functoriality, say we have two chain maps

$$\mathbf{A} \xrightarrow{f} \mathbf{B} \xrightarrow{g} \mathbf{C}. \quad (24.2.12)$$

We need to show the commutativity of the triangle

$$\begin{array}{ccc} H_n(\mathbf{A}) & \xrightarrow{H_n(f)} & H_n(\mathbf{B}) \\ & \searrow H_n(g \circ f) & \swarrow H_n(g) \\ & H_n(\mathbf{C}) & \end{array} \quad (24.2.13)$$

We have

$$\begin{aligned} H_n(g \circ f)([z]) &= [(g \circ f)_n(z)] = [g_n \circ f_n(z)] = H_n(g)([f_n(z)]) = \\ &= H_n(g)(H_n(f)[z]) = H_n(g)H_n(f)([z]). \end{aligned} \quad (24.2.14)$$

Finally, additivity is trivial since f_n are homomorphisms:

$$H_n(f + f') = H_n(f) + H_n(f'). \quad (24.2.15)$$

■

Remark 24.2.7. On cochain complexes, this functor is of course contravariant.

From now on we write

$$f_* = H_n(f), \quad g^* = H^n(g) \quad (24.2.16)$$

for the descendants of maps in (co)homology.

Lemma 24.2.8. *Exact additive covariant functors preserve homology: $H_*(F(\mathbf{C})) \cong F(H_*(\mathbf{C}))$.*

Proof. This follows from the fact that exact additive functors preserve kernels and quotients (and all functors preserve images). ■

Proposition 24.2.9. *The direct limit of the homology groups of a directed system of chain complexes \mathbf{C}^i is the homology group of the direct limit:*

$$\varinjlim_i H_n(\mathbf{C}^i) \cong H_n(\varinjlim_i \mathbf{C}^i). \quad (24.2.17)$$

Proof. This follows immediately from the preceding Lemma and the fact that direct limits preserve exactness in abelian categories (Proposition 22.2.6). ■

Remark 24.2.10. Note that the analogous statement for homotopy groups π_n was proven by us only for sequences of CW subcomplexes (Corollary 5.3.8) and does not hold in general even for π_1 .

Theorem 24.2.11 (Connecting homomorphism lemma). *Let*

$$0 \rightarrow \mathbf{A} \xrightarrow{i} \mathbf{B} \xrightarrow{\pi} \mathbf{C} \rightarrow 0 \quad (24.2.18)$$

be a short exact sequence of chain complexes. Then for each n there is a connecting homomorphism

$$\delta = \delta_n : H_n(\mathbf{C}) \rightarrow H_{n-1}(\mathbf{A}), \quad [z] \rightarrow [i_{n-1}^{-1} d_n \pi_n^{-1}(z)]. \quad (24.2.19)$$

Proof. Several checks need to be done. First, why does $d_n \pi_n^{-1}(z) \in \text{im } i_{n-1}$? Let $y : \pi_n(y) = z$, then $d_n(y) \in B_{n-1}$. Since $z \in Z_n(\mathbf{C})$ we have $\pi_{n-1}(d_n(y)) = d_n(\pi_n(y)) = d_n(z) = 0$. Therefore $\exists! x \in A_{n-1} : i_{n-1}(x) = d_n(y)$, and we have a well defined map

$$z \mapsto x. \quad (24.2.20)$$

Next, we need to verify that $d_{n-1}(x) = 0$ and the class $[x] \in H_{n-1}(\mathbf{A})$ is independent of the choice of y . Moreover, one needs to check that if $[z] = [z']$, then $[x] = [x']$. Lastly, we need to check that δ is a homomorphism. These checks are left as an exercise. ■

Theorem 24.2.12 (Zig-zag lemma/Long exact sequence in homology). *For any short exact sequence of complexes*

$$0 \rightarrow \mathbf{A} \xrightarrow{i} \mathbf{B} \xrightarrow{\pi} \mathbf{C} \rightarrow 0, \quad (24.2.21)$$

the induced sequence in homology

$$\cdots \rightarrow H_{n+1}(\mathbf{C}) \xrightarrow{\delta_{n+1}} H_n(\mathbf{A}) \xrightarrow{i_*} H_n(\mathbf{B}) \xrightarrow{\pi_*} H_n(\mathbf{C}) \xrightarrow{\delta_n} H_{n-1}(\mathbf{A}) \rightarrow \cdots \quad (24.2.22)$$

is exact.

Proof. It is easy to see that for any complex \mathbf{C} the two “fundamental sequences”

$$0 \rightarrow Z_n(\mathbf{C}) \rightarrow C_n \xrightarrow{d_n} B_{n-1}(\mathbf{C}) \rightarrow 0, \quad (24.2.23)$$

$$0 \rightarrow B_n(\mathbf{C}) \rightarrow Z_n(\mathbf{C}) \rightarrow H_n(\mathbf{C}) \rightarrow 0 \quad (24.2.24)$$

are short exact. This one is also obviously exact:

$$0 \rightarrow B_n(\mathbf{C}) \rightarrow C_n \rightarrow C_n/B_n(\mathbf{C}) \rightarrow 0. \quad (24.2.25)$$

Using the exactness of the first sequence, we also have the exactness of this one:

$$0 \rightarrow H_n(\mathbf{C}) \rightarrow C_n/B_n(\mathbf{C}) \rightarrow \underbrace{B_{n-1}(\mathbf{C})}_{\cong C_n/Z_n(\mathbf{C})} \rightarrow 0. \quad (24.2.26)$$

Putting it all together, we conclude the exactness of this sequence:

$$0 \rightarrow H_n(\mathbf{C}) \rightarrow C_n/B_n(\mathbf{C}) \rightarrow Z_{n-1}(\mathbf{C}) \rightarrow H_{n-1}(\mathbf{C}) \rightarrow 0. \quad (24.2.27)$$

Now let us use this sequence for the columns of one large diagram:

$$\begin{array}{ccccccc}
 & 0 & & 0 & & 0 & \\
 & \downarrow & & \downarrow & & \downarrow & \\
 H_n(\mathbf{A}) & \longrightarrow & H_n(\mathbf{B}) & \longrightarrow & H_n(\mathbf{C}) & \dashrightarrow & \\
 & \downarrow & & \downarrow & & \downarrow & \\
 A_n/B_n(\mathbf{A}) & \longrightarrow & B_n/B_n(\mathbf{B}) & \longrightarrow & C_n/B_n(\mathbf{C}) & \longrightarrow & 0 \\
 & | & | & | & | & | & \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
 0 & \longrightarrow & Z_{n-1}(\mathbf{A}) & \longrightarrow & Z_{n-1}(\mathbf{B}) & \longrightarrow & Z_{n-1}(\mathbf{C}) \\
 & \downarrow & & \downarrow & & \downarrow & \\
 & \dashrightarrow & H_{n-1}(\mathbf{A}) & \longrightarrow & H_{n-1}(\mathbf{B}) & \longrightarrow & H_{n-1}(\mathbf{C}) \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
 \end{array}$$

(24.2.28)

The exactness of the rows can be checked directly using the exactness of $\mathbf{A} \rightarrow \mathbf{B} \rightarrow \mathbf{C}$. Then this is in fact exactly the type of diagram for which we can apply the Snake Lemma 22.3.8 and establish the existence of δ_n . It only remains to check that the construction of δ_n in the Snake Lemma exactly coincides with the one in Proposition 24.2.11. ■

Corollary 24.2.13. *The Snake lemma is equivalent to the Zig-zag lemma.*

Proof. The proof of Theorem 24.2.12 shows that the Snake Lemma implies the long exact sequence in homology. The converse also trivially holds since we can apply the Zig-zag lemma to a short exact sequence of complexes concentrated in just two degrees of the form $0 \rightarrow \ker \pi \rightarrow \operatorname{im} \pi \rightarrow 0$ (let the two other complexes have ρ and σ instead of π) and obtain the long exact sequence of homologies, which reads

$$0 \rightarrow \ker \pi \rightarrow \ker \rho \rightarrow \ker \sigma \rightarrow \operatorname{coker} \pi \rightarrow \operatorname{coker} \rho \rightarrow \operatorname{coker} \sigma \rightarrow 0, \quad (24.2.29)$$

which is exactly the Snake Lemma 22.3.8. ■

Proposition 24.2.14 (Naturality of the connecting homomorphism³⁶). *Given a commutative*

diagram with exact rows in the category $\text{Comp}(C)$

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathbf{A} & \xrightarrow{i} & \mathbf{B} & \xrightarrow{p} & \mathbf{C} & \longrightarrow 0 \\ & & f \downarrow & & g \downarrow & & h \downarrow & \\ 0 & \longrightarrow & \mathbf{A}' & \xrightarrow{j} & \mathbf{B}' & \xrightarrow{q} & \mathbf{C}' & \longrightarrow 0 \end{array} \quad (24.2.30)$$

there is a commutative diagram in C with exact rows

$$\begin{array}{ccccccc} \longrightarrow H_n(\mathbf{A}) & \xrightarrow{i_*} & H_n(\mathbf{B}) & \xrightarrow{p_*} & H_n(\mathbf{C}) & \xrightarrow{\delta} & H_{n-1}(\mathbf{A}) & \longrightarrow \\ \downarrow f_* & & \downarrow g_* & & \downarrow h_* & & \downarrow f_* & \\ \longrightarrow H_n(\mathbf{A}') & \xrightarrow{j_*} & H_n(\mathbf{B}') & \xrightarrow{q_*} & H_n(\mathbf{C}') & \xrightarrow{\delta'} & H_{n-1}(\mathbf{A}') & \longrightarrow \end{array} \quad (24.2.31)$$

In other words, morphisms between exact sequences of complexes naturally induce morphisms between the exact sequences in homology, i.e. the connecting homomorphism δ is natural w.r.t. such morphisms.

Proof. The exactness of the rows is the content of the theorem about the long exact sequence. The commutativity of the first two squares follows from the functoriality of H_n . Checking the commutativity of the square $\delta' \circ h_* = f_* \circ \delta$ requires expanding the first commutative diagram in $\text{Comp}(C)$ into a 3-dimensional diagram in C .

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A_n & \xrightarrow{i} & B_n & \xrightarrow{p} & C_n & \longrightarrow & 0 \\ & & \downarrow \delta & & \downarrow f_* & & \downarrow \delta & & \\ 0 & \longrightarrow & A_{n-1} & \xrightarrow{i} & B_{n-1} & \xrightarrow{p} & C_{n-1} & \longrightarrow & 0 \\ & & \downarrow & & \downarrow g_* & & \downarrow h_* & & \\ 0 & \longrightarrow & A'_n & \xrightarrow{j} & B'_n & \xrightarrow{q} & C'_n & \longrightarrow & 0 \\ & & \downarrow f_* & & \downarrow g_* & & \downarrow h_* & & \\ 0 & \longrightarrow & A'_{n-1} & \xrightarrow{j} & B'_{n-1} & \xrightarrow{q} & C'_n & \longrightarrow & 0 \end{array} \quad (24.2.32)$$

Taking $[c] \in H_n(\mathbf{C})$, we must show that $f_*\delta[c] = \delta'h_*[c]$. Let $b \in B_n$ be such that $p(b) = c$. Then $\delta[c] = [a]$, where $i(a) = d(b)$. Hence $f_*\delta[c] = [f(a)]$.

On the other hand, since h is a chain map, $q(g(b)) = h(p(b)) = h(c)$. Hence we can use $g(b)$ as the pre-image of $h(c)$ in B'_n , so by construction of the connecting homomorphism, $\delta'[h(c)] = [a']$ where $j(a') = d(g(b))$. But $j(f(a)) = g(i(a)) = g(d(b)) = d(g(b)) =$

$j(a')$ and so $f(a) = a'$ because j is injective. Therefore $\delta' h_*[c] = [a'] = [f(a)]$, which coincides with the l.h.s. computed above. ■

Naturality is instrumental in many arguments involving long exact sequences because it means that we can apply sequences of chain maps to long exact sequences and preserve the commutativity of diagrams.

Proposition 24.2.15 (Algebraic Mayer-Vietoris sequence). *If in the setting of Theorem 24.2.14, h_* is also an isomorphism in homology, then one has the long exact sequence*

$$\cdots \rightarrow H_{n+1}(\mathbf{B}') \xrightarrow{\Delta} H_n(\mathbf{A}) \xrightarrow{(f_*, i_*)} H_n(\mathbf{A}') \oplus H_n(\mathbf{B}) \xrightarrow{j_* - g_*} H_n(\mathbf{B}') \xrightarrow{\Delta} H_{n-1}(\mathbf{A}) \rightarrow \cdots \quad (24.2.33)$$

where $\Delta = \delta \circ h_*^{-1} \circ q_*$.

Proof. This is an exercise in diagram chasing. ■

The relation of this Proposition to the **MV** sequence in de Rham cohomology will become clear when we get to relative (co)homology and the excision property. The isomorphism h_* in question is $H_n(U, U \cap V) \cong H_n(M, V)$ for $M = U \cup V$. A rough visualization of $H_n(X, Y)$ with $Y \subset X$ is the homology of the space obtained from X by contracting Y into one point. Then it is clear that contracting V inside M has the “same” result as contracting $U \cap V$ inside U .

Definition 24.2.16 (Maps of degree p). Let \mathbf{C} and \mathbf{D} be two chain complexes in C . A map of degree p between them, denoted $s : \mathbf{C} \rightarrow \mathbf{D}$, is a collection of maps $s_n : C_n \rightarrow D_{n+p}$. Note that no commutativity is required here!

Now that we know that H_n is a functor, it is natural to ask how much $H_n(f)$ actually depends on f . It turns out that $H_n(f)$ is invariant under homotopies of f defined in the following categorical sense.

Definition 24.2.17. Let $f, g : \mathbf{B} \rightarrow \mathbf{C}$ be two chain maps. They are called homotopic ($f \sim g$) if there is a map of degree +1 denoted $s : \mathbf{B} \rightarrow \mathbf{C}$ such that

$$\forall n, \quad f_n - g_n = d_{n+1}s_n + s_{n-1}d_n. \quad (24.2.34)$$

(As usual, we use d_n to denote the differentials in all complexes at once). This can be

visualized by the following *non-commutative* diagram:

$$\begin{array}{ccccccc} \cdots & \longrightarrow & B_{n+1} & \longrightarrow & B_n & \longrightarrow & B_{n-1} \longrightarrow \cdots \\ & & f_{n+1} \downarrow & \swarrow g_{n+1} & f_n \downarrow & \swarrow g_n & f_{n-1} \downarrow \swarrow g_{n-1} \\ \cdots & \longrightarrow & C_{n+1} & \longrightarrow & C_n & \longrightarrow & C_{n-1} \longrightarrow \cdots \end{array} \quad (24.2.35)$$

Theorem 24.2.18. *Homotopic chain maps induce the same morphism in homology.*

Proof. Let $s : \mathbf{B} \rightarrow \mathbf{C}$ be the homotopy. If z is an n -cycle, dropping subscripts, $dz = 0$, then

$$f(z) - g(z) = d(s(z)) + s(d(z)) = d(s(z)), \quad (24.2.36)$$

i.e. $f(z) - g(z) \in B_n(\mathbf{C})$ and $f_* = g_*$. ■

Definition 24.2.19 (Acyclic, contractible complexes). A complex is called acyclic if all its (co)homologies are trivial. A complex \mathbf{C} is called contractible if there exists a homotopy $1_{\mathbf{C}} \sim 0_{\mathbf{C}}$. All contractible complexes are acyclic since contractibility implies $1_{H_n(\mathbf{C})} = 0_{H_n(\mathbf{C})}$.

Definition 24.2.20 (Quasi-isomorphisms). Two complexes \mathbf{A}, \mathbf{B} are called quasi-isomorphic (or *weakly equivalent*) if there is a chain map $f : \mathbf{A} \rightarrow \mathbf{B}$ such that $f_* = H_n(f) : H_n(\mathbf{A}) \rightarrow H_n(\mathbf{B})$ is an isomorphism for all n .

Definition 24.2.21 (Homotopy equivalence). Two complexes \mathbf{A}, \mathbf{B} are called homotopy equivalent if there are two chain maps $f : \mathbf{A} \rightarrow \mathbf{B}$ and $g : \mathbf{B} \rightarrow \mathbf{A}$ such that

$$f \circ g \sim 1_{\mathbf{B}}, \quad g \circ f \sim 1_{\mathbf{A}}. \quad (24.2.37)$$

Clearly homotopy equivalent complexes are quasi-isomorphic.

Exercise 24.5. Check that a Lie derivative \mathcal{L}_X is a cochain map on the de Rham complex. Furthermore, show that applying a Lie derivative \mathcal{L}_X to a differential form doesn't change its class in de Rham cohomology. For this, find the homotopy operator between \mathcal{L}_X and the zero morphism.

24.3 Singular homology

de Rham cohomology is defined only for smooth manifolds, and simplicial homology is defined only for triangulable spaces. The type of homology we define now works for all topological spaces, but proving its properties for nicer spaces (like manifolds) requires extra

work. The idea is, instead of trying to break up our topological space into simplices, to study the set of all continuous maps from simplices to our space.

Definition 24.3.1 (Singular simplices). We denote by Δ^n the standard n -simplex $\Delta^n = \{\sum_{i=0}^n \alpha_i e_i \mid \sum_i \alpha_i = 1\}$, where $\{e_i\}_{i=0}^n$ is the standard basis in \mathbb{R}^{n+1} . Let X be a topological space. A singular n -simplex in X is a continuous map $\sigma \in C(\Delta^n, X)$ (no other constraints, hence “singular”). The group $C_n(X)$ of singular n -chains is the free abelian group generated by the set $C(\Delta^n, X)$. Sometimes we will write $C_n^{\text{sing}}(X)$ to distinguish from other kinds of chain groups.

Definition 24.3.2 (Singular homology). The boundary map $\partial_n : C_n(X) \rightarrow C_{n-1}(X)$ is defined as before:

$$\partial_n \sigma = \sum_i (-1)^i \sigma|_{\langle e_0, \dots, \hat{e}_i, \dots, e_n \rangle}, \quad (24.3.1)$$

where we implicitly identify the faces of Δ^n with Δ^{n-1} while preserving the order of vertices, so that $\sigma|_{\langle e_0, \dots, \hat{e}_i, \dots, e_n \rangle}$ is a singular $(n-1)$ -simplex.

Singular homology groups are defined as

$$H_n(X) = \ker \partial_n / \text{im } \partial_{n+1}. \quad (24.3.2)$$

Unlike with simplicial homology, here it is immediately obvious that homeomorphic spaces have isomorphic singular homologies. The price we pay for this is that the rank of $C_n^{\text{sing}}(X)$ is uncountably large, so it is not clear at a glance whether singular homologies of spaces that have finitely generated simplicial homologies are also finitely generated.

Proposition 24.3.3. *If X consists of l path-connected components, then $H_0(X) \cong \mathbb{Z}^l$.*

Proof. It suffices to prove this for $l = 1$, i.e. a path-connected X . We have $H_0(X) = C_0(X)/\text{im } \partial_1$. A singular 0-chain in X is just a finite sum $\sum_i n_i p_i$ where p_i are points in X and $n_i \in \mathbb{Z}$. Define the homomorphism $\epsilon : C_0(X) \rightarrow \mathbb{Z}$, called the *augmentation map*, by

$$\epsilon \left(\sum_i n_i p_i \right) = \sum_i n_i. \quad (24.3.3)$$

It is surjective as long as $X \neq \emptyset$. The claim is that $\ker \epsilon \cong \text{im } \partial_1$ if X is path-connected, so ϵ induces an isomorphism $H_0(X) \cong \mathbb{Z}$.

First, it is clear by definition of ∂ that $\text{im } \partial_1 \subset \ker \epsilon$. For the reverse inclusion, suppose $\epsilon(\sum_i n_i p_i) = 0$. Choose a path γ_i connecting a base point x_0 to p_i . We can view γ_i as a singular 1-simplex and have $\partial \gamma_i = p_i - x_0$. Hence $\partial(\sum_i n_i \gamma_i) = \sum_i n_i p_i$ since $\sum_i n_i = 0$. Therefore $\sum_i n_i p_i$ is a boundary, which shows $\ker \epsilon \subset \text{im } \partial_1$. ■

Proposition 24.3.4. If X consists of a single point, then $H_n(X) = 0$ for $n > 0$ and $H_0(X) \cong \mathbb{Z}$.

Proof. In this case there is a unique singular n -simplex σ_n for each n and

$$\partial\sigma_n = \sum_{i=0}^n (-1)^i \sigma_{n-1} = \begin{cases} 0, & n \text{ odd}, \\ \sigma_{n-1}, & n \text{ even}. \end{cases} \quad (24.3.4)$$

This the singular chain complex reads

$$\cdots \rightarrow \mathbb{Z} \xrightarrow{\cong} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{\cong} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \rightarrow 0. \quad (24.3.5)$$

The homology of this complex is trivial except for $H_0 \cong \mathbb{Z}$. ■

Sometimes it is very helpful to work with a slightly modified homology that vanishes entirely for a point. The following definition achieves this.

Definition 24.3.5 (Reduced homology). The reduced homology groups $\tilde{H}_n(X)$ are the homology groups of the *augmented chain complex*

$$\cdots C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\epsilon} \mathbb{Z} \rightarrow 0. \quad (24.3.6)$$

The reduced homology of a point is trivial in all degrees. Also $H_n(X) \cong \tilde{H}_n(X)$ for all $n > 0$ and $H_0(X) = \tilde{H}_0(X) \oplus \mathbb{Z}$ since ϵ induces a map $H_0(X) \rightarrow \mathbb{Z}$ with kernel $\tilde{H}_0(X)$.

24.4 Homotopy invariance

Here we show that singular homology is homotopy invariant.

Proposition 24.4.1 ([22, Thm. 2.10]). If two maps $f, g \in C(X, Y)$ between topological spaces X, Y are homotopic, then they induce the same homomorphism in homology

$$f_* = g_* : H_n(X) \rightarrow H_n(Y) \quad (24.4.1)$$

for all n . In particular, homotopy equivalent spaces have isomorphic homology groups.

Proof. The key is learning to subdivide the prism $\Delta^n \times I$ (where $I = [0, 1]$) into simplices. If $\langle v_0, \dots, v_n \rangle$ is the copy of Δ^n at the bottom of the prism and $\langle w_0, \dots, w_n \rangle$ is the one at the top, then it is easy to check that the prism is triangulated by the $(n+1)$ -simplices $\langle v_0, \dots, v_i, w_i, \dots, w_n \rangle$, each intersecting the next in an n -simplex face. The cases $n = 1, 2$ are shown in figure 24.2.

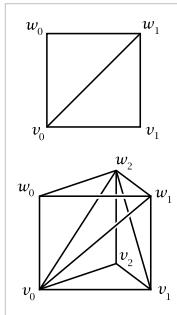


Fig. 24.2: Triangulation of a prism, see Proposition 24.4.1.

Given a homotopy $F : X \times I \rightarrow Y$ from f to g and a singular simplex $\sigma : \Delta^n \rightarrow X$, we form the composition $F \circ (\sigma \times \text{id}) : \Delta^n \times I \rightarrow X \times I \rightarrow Y$. Using this, we define the *prism operators*

$$P : C_n(X) \rightarrow C_{n+1}(Y), \quad P(\sigma) = \sum_i (-1)^i F \circ (\sigma \times \text{id}) \langle x_0, \dots, v_i, w_i, \dots, w_n \rangle. \quad (24.4.2)$$

We claim that P is a homotopy operator between these two chain complexes, namely

$$g_* - f_* = \partial P + P\partial. \quad (24.4.3)$$

The geometric meaning of $\partial P = g_* - f_* - P\partial$ is that the boundary of the prism equals its top $\Delta^n \times \{1\}$, minus its bottom $\Delta^n \times \{0\}$, and plus the sides $\partial \Delta^n \times I$ (compare this with Cartan's magic formula and its cohomological meaning, which is nothing but a dual to this geometric statement, see Exercise 24.5).

The verification of this identity follows easily by direct expansion and use of the definition of the homotopy F and $g \circ \sigma = g_*(\sigma)$ (see [22, Thm. 2.10] for full details). ■

Corollary 24.4.2. *If X is contractible, then $\tilde{H}_n(X) = 0$ for any n .*

Continuous maps induce homomorphisms in reduced homology as well since the augmentation map ϵ commutes with f_* . Then the same theorem holds for reduced homologies with the same proof.

24.5 Relative homology and excision

Relative homology generalizes the idea behind the **MV** sequence in de Rham cohomology. Given a subset $A \subset X$ in a topological space, we ask how similar the groups $H_n(X/A)$ and $H_n(X)/H_n(A)$ are. The difference, in a sense, is measured by the relative homology.

Definition 24.5.1. Let X be a topological space and $A \subset X$ a topological subspace (i.e. a subset with the induced subset topology). The group of singular chains $C_n(A)$ can be treated as a subgroup of $C_n(X)$. Define the relative singular chain groups as $C_n(X, A) = C_n(X)/C_n(A)$. The boundary operator ∂ descends to these groups because $\partial(C_n(A)) \subset C_{n-1}(A)$. The homology of the resulting chain complex is called the relative homology of the pair (X, A) .

Relative homology is a covariant functor on the category of topological pairs, $H_n : \text{TopPair} \rightarrow \text{Ab}$.

Proposition 24.5.2 (Exact sequence of a pair). *The exact sequence of relative chain groups*

$$0 \rightarrow C_\bullet(A) \rightarrow C_\bullet(X) \rightarrow C_\bullet(X, A) \rightarrow 0 \quad (24.5.1)$$

induces a long exact sequence in relative homology

$$\cdots \rightarrow H_n(A) \rightarrow H_n(X) \rightarrow H_n(X, A) \rightarrow H_{n-1}(A) \rightarrow H_{n-1}(X) \rightarrow H_{n-1}(X, A) \rightarrow \cdots \quad (24.5.2)$$

Proof. The first exact sequence is the definition of relative homology and the second one is an application of the Zig-zag lemma. ■

The same sequence exists in reduced homology. This exact sequence has a natural geometric interpretation: elements of $H_n(X, A)$ can be thought of as n -chains in X that differ from an actual cycle in X only by a chain in A , and the connecting homomorphism maps such a cycle into an $(n - 1)$ -cycle in A by simply computing the boundary.

Corollary 24.5.3.

1. $H_n(X, \{x_0\}) \cong \tilde{H}_n(X)$ for all n .
2. If A is contractible, then $H_n(X, A) \cong \tilde{H}_n(X)$ for all n ;
3. if X is contractible, then $H_{n+1}(X, A) \cong \tilde{H}_n(A)$ for all n .

Example 24.5.4 (Homology of a bouquet of circles). Consider $X = \mathbb{R}$ and $A \subset \mathbb{R}$ a finite subset of size $(k + 1)$. Then X/A is homotopy equivalent to a wedge sum of k circles (a.k.a. a bouquet of circles). We will later show that $H_1(X, A) \cong \tilde{H}_1(X/A) = H_1(X/A)$, so we have $H_1(\bigvee_{i=1}^k \mathbb{S}^1) = \tilde{H}_0(A) = \mathbb{Z}^k$. We also see that all higher homologies of the bouquet vanish. Another way to obtain this result will be from the Hurewicz theorem.

Proposition 24.5.5 (Exact sequence of a triple). *For a triple (X, A, B) where $B \subset A \subset X$, the exact sequence of relative chain complexes*

$$0 \rightarrow C_*(A, B) \rightarrow C_*(X, B) \rightarrow C_*(X, A) \rightarrow 0 \quad (24.5.3)$$

induces a long exact sequence in relative homology

$$\cdots \rightarrow H_n(A, B) \rightarrow H_n(X, B) \rightarrow H_n(X, A) \rightarrow H_{n-1}(A, B) \rightarrow H_{n-1}(X, B) \rightarrow H_{n-1}(X, A) \rightarrow \cdots \quad (24.5.4)$$

Proof. Exercise. ■

Proposition 24.5.6 (Homotopy property for relative homology). *If two maps $f, g : (X, A) \rightarrow (Y, B)$ in the category of topological pairs are homotopic also through maps of pairs, then they induce the same homomorphism $f_* = g_* : H_n(X, A) \rightarrow H_n(Y, B)$.*

Proof. The prism operator constructed in the proof of the homotopy property induces a relative prism operator $P : C_n(X, A) \rightarrow C_{n+1}(Y, B)$. Since we are passing to quotient groups, the formula $\partial P + P\partial = g_* - f_*$ still holds and f_*, g_* are chain homotopic on relative chain groups. ■

For a space X , let $\mathcal{U} = \{U_\alpha\}_\alpha$ be a collection of subspaces of X whose interiors cover X and let $C_n^{\mathcal{U}}(X)$ be the subgroup of $C_n(X)$ generated by simplices whose images lie inside one of the sets in the cover. The boundary operator ∂ maps $C_n^{\mathcal{U}}(X)$ to $C_{n-1}^{\mathcal{U}}(X)$ and therefore we have a chain complex $C_{\bullet}^{\mathcal{U}}(X)$ with homology groups denoted $H_n^{\mathcal{U}}(X)$.

Proposition 24.5.7 ([22, Thm. 2.21]). *With the notation just described, the inclusion $i : C_n^{\mathcal{U}}(X) \hookrightarrow C_n(X)$ is a chain homotopy equivalence and hence induces isomorphisms*

$$H_n^{\mathcal{U}}(X) \cong H_n(X), \quad n \geq 0. \quad (24.5.5)$$

Proof. The proof is fairly tedious with multiple technical steps and we refer the reader to [22, Thm. 2.21] for all the details. The idea is to perform iterated barycentric subdivision of all simplices to construct a quasi-inverse (chain homotopy inverse) to i .

1. *Barycentric subdivision of simplices.* Denote the barycenter of an n -simplex $\sigma = \langle v_0, \dots, v_n \rangle$ by $b = \sum_i v_i / (n+1)$. The barycentric subdivision consists of n -simplices $\langle b, w_0, \dots, w_{n-1} \rangle$ where iteratively $\langle w_0, \dots, w_{n-1} \rangle$ is an $(n-1)$ -simplex in the barycentric subdivision of a face $\langle v_0, \dots, \hat{v}_i, \dots, v_n \rangle$.

To us it is important that the diameter of any simplex in the barycentric subdivision of σ is bounded by $\frac{n}{n+1} \cdot \text{diam}(\sigma)$. Since $\frac{n}{n+1} < 1$, iterating the subdivision creates simplices uniformly and arbitrarily small in size.

2. *Barycentric subdivision of linear chains.* For a convex set Y in a Euclidean space, all linear maps $\Delta^n \rightarrow Y$ generate a subgroup of $C_n(Y)$ that we denote by $L_n(Y)$, the linear chains. They form a chain complex with the same boundary operator. Each such linear simplex can be identified with a regular simplex in Y itself. For convenience we augment the complex with $L_{-1}(Y) = \mathbb{Z}$ generated by the empty simplex $\langle \emptyset \rangle$ with $\partial\sigma = \langle \emptyset \rangle$ for all $\sigma \in L_1(Y)$.

Each point $b \in Y$ determines a homomorphism $b : L_n(Y) \rightarrow L_{n+1}(Y)$ defined on simplices by $b(\langle w_0, \dots, w_n \rangle) = \langle b, w_0, \dots, w_n \rangle$ (a “cone operator”). It is easy to check that $\partial b + b\partial = \text{id}$, so b is a chain homotopy between the identity and zero on the augmented chain complex $L_\bullet(Y)$.

Define the subdivision homomorphism $S : L_n(Y) \rightarrow L_n(Y)$ by induction in the degree. Denoting by b_λ the image of the barycenter in a linear singular simplex $\lambda : \Delta^n \rightarrow Y$. Then the inductive formula for S is $S(\lambda) = b_\lambda(S\partial\lambda)$, where b_λ was defined above.

Then one checks that $\partial S = S\partial$ so that S provides a chain map from $L_\bullet(Y)$ to itself. Next we build a chain homotopy $T : L_n(Y) \rightarrow L_{n+1}(Y)$ between S and the identity. It is defined inductively by $T_{n=-1} = 0$ and $T(\lambda) = b_\lambda(\lambda - T\partial\lambda)$ for $n \geq 0$. The geometric interpretation of this formula is that we inductively subdivide $\Delta^n \times I$ by joining all simplices in the bottom and side faces of the prism to the barycenter of the top face, and T takes the image of this subdivision under the projection $\Delta^n \times I \rightarrow \Delta^n$. Finally one verifies that $\partial T + T\partial = \text{id} - S$.

3. *Barycentric subdivision of general chains.* Define $S : C_n(X) \rightarrow C_n(X)$ by setting $S\sigma = \sigma_*S(\Delta^n)$. It follows easily that it is a chain map, $\partial S = S\partial$. Then in a similar fashion we define $T : C_n(X) \rightarrow C_{n+1}(X)$ by $T(\sigma) = \sigma_*T(\Delta^n)$, and we also have $\partial T + T\partial = \text{id} - S$.
4. *Iterated barycentric subdivision.* We can construct a chain homotopy between id and the iterated subdivision \mathbb{S}^m , given by $D_m = \sum_{0 \leq i < m} T\mathbb{S}^i$ (easy check that $\partial D_m + D_m\partial = \text{id} - \mathbb{S}^m$).

For any given simplex σ there exists a number $m(\sigma)$ such that $\mathbb{S}^{m(\sigma)}(\sigma)$ lies in $C_n^U(X)$ (because the diameter of the simplices in $\mathbb{S}^m(\Delta^n)$ will be less than the strictly positive Lebesgue number of the cover of the compact metric space Δ^n by the open sets $\sigma^{-1}(\text{int } U_\alpha)$ for large m , and by definition a set of diameter less than the Lebesgue number is contained in one of the sets in the cover). Now we define $D : C_n(X) \rightarrow C_{n+1}(X)$ by $D(\sigma) = D_{m(\sigma)}(\sigma)$. For this operator we need to find a chain map $\rho : C_n(X) \rightarrow C_n(X)$ such that $\partial D + D\partial = \text{id} - \rho$. This formula itself hints that $\rho(\sigma) = \mathbb{S}^{m(\sigma)}(\sigma) +$

$D_{m(\sigma)}(\partial\sigma) - D(\partial\sigma)$ does the job. Viewing ρ as a map $C_n(X) \rightarrow C_n^{\mathcal{U}}(X)$, we have $\partial D + D\partial = \text{id} - i \circ \rho$. We also know $\rho \circ i = \text{id}$ since D is identically zero and $m(\sigma) = 0$ on $C_n^{\mathcal{U}}(X)$. Therefore ρ is a chain homotopy inverse for i . ■

Theorem 24.5.8 (Excision in homology). *Given subspaces $Z \subset A \subset X$ such that $\bar{Z} \subset \text{int } A$, the inclusion $(X \setminus Z, A \setminus Z) \hookrightarrow (X, A)$ induces isomorphisms*

$$H_n(X \setminus Z, A \setminus Z) \cong H_n(X, A) \quad (24.5.6)$$

for all n . Equivalently, for subspaces $A, B \subset X$ that cover X , the inclusion $(B, A \cap B) \hookrightarrow (X, A)$ induces isomorphisms

$$H_n(B, A \cap B) \cong H_n(X, A) \quad (24.5.7)$$

for all n (by setting $B = X \setminus Z$, $Z = X \setminus B$, in which case $A \cap B = A \setminus Z$ and the condition $\bar{Z} \subset \text{int } A$ is equivalent to $X = \text{int } A \cup \text{int } B$ since $X - \text{int } B = \bar{Z}$).

Proof. We prove the version with $X = A \cup B$. For the cover $\mathcal{U} = \{A, B\}$ we have the chain groups $C_n^{\mathcal{U}}(X)$ consisting of sums of chains in A and chains in B . At the end of the preceding proof we established $\partial D + D\partial = \text{id} - i \circ \rho$ and $\rho \circ i = \text{id}$. All maps here take chains in A to chains in A , so they induce quotient maps when we factor out chains in A . These quotient maps automatically satisfy the same formulas, so the inclusion $C_n^{\mathcal{U}}(X)/C_n(A) \hookrightarrow C_n(X)/C_n(A)$ induces an isomorphism on homology. The map $C_n(B)/C_n(A \cap B) \rightarrow C_n^{\mathcal{U}}(X)/C_n(A)$ induced by the inclusion is obviously an isomorphism since both quotient groups are free generated by the singular n -simplices in B that do not lie in A . Hence we obtain the desired isomorphism $H_n(B, A \cap B) \cong H_n(X, A)$, induced by inclusion. ■

Corollary 24.5.9 (Mayer-Vietoris Sequence). *If $X = U \cup V$ where U, V are open, then the Mayer-Vietoris sequence in singular homology is exact:*

$$\cdots \rightarrow H_n(U \cap V) \xrightarrow{j_{U*} \oplus j_{V*}} H_n(U) \oplus H_n(V) \xrightarrow{i_{U*} - i_{V*}} H_n(X) \xrightarrow{\Delta_n} H_{n-1}(A \cap V) \rightarrow \cdots . \quad (24.5.8)$$

Proof. We have the commutative square of inclusions

$$\begin{array}{ccc} U \cap V & \xrightarrow{j_U} & U \\ j_V \downarrow & & \downarrow i_U \\ V & \xrightarrow{i_V} & X. \end{array} \quad (24.5.9)$$

The naturality of the long exact sequence in homology with respect to morphisms of pairs implies the following commutative ladder with exact rows:

$$\begin{array}{ccccccc} \rightarrow H_{n+1}(U, U \cap V) & \xrightarrow{\delta_{n+1}} & H_n(U \cap V) & \xrightarrow{j_{U*}} & H_n(U) & \rightarrow H_n(U, U \cap V) & \rightarrow \\ \cong \downarrow \iota_* & & \downarrow j_{V*} & & \downarrow i_{U*} & & \cong \downarrow \iota_* \\ \longrightarrow H_{n+1}(X, V) & \xrightarrow{\delta'_{n+1}} & H_n(V) & \xrightarrow{i_{V*}} & H_n(X) & \longrightarrow H_n(X, V) & \longrightarrow \end{array} \quad (24.5.10)$$

Since we are in the situation of the Excision Theorem 24.5.8, all the induced maps ι_* are isomorphisms. Thus, if we denote by $\Delta_n : H_n(X) \rightarrow H_{n-1}(U \cap V)$ the homomorphism

$$\Delta_n : H_n(X) \rightarrow H_n(X, V) \xrightarrow{\iota_*^{-1}} H_n(U, U \cap V) \rightarrow H_{n-1}(U \cap V), \quad (24.5.11)$$

then the algebraic MV sequence (Theorem 24.2.15) produces exactly the asserted sequence. ■

The MV sequence, and excision more generally, allows for inductive computations of homology groups. For example, one can easily compute the homology groups of spheres in a way similar to how we did it using excision in homotopy theory. However, the same result will follow more easily from other general theorems we will prove in the following sections.

Exercise 24.6. Compute the singular homology groups of the sphere \mathbb{S}^n using the MV sequence.

Definition 24.5.10 (Good pair). (X, A) is called a good pair if A is a deformation retract of an open subset in X . That is, there is an open set U such that $A \subset U \subset X$ and a map $F \in C(U \times [0, 1], X)$ such that $F(x, 0) = x$, $F(x, 1) \in A$ for all $x \in U$ and also $F(x, t) = x$ for all $x \in A$ and $t \in [0, 1]$.

Proposition 24.5.11. *For good pairs (X, A) the quotient map $q : (X, A) \rightarrow (X/A, A/A)$ induces isomorphisms*

$$H_n(X, A) \xrightarrow{q_*} H_n(X/A, A/A) \cong \tilde{H}_n(X/A), \quad n \geq 0. \quad (24.5.12)$$

Proof. Let U be a neighborhood of A in X that deformation retracts onto A . We have the commutative diagram (the commutativity is easy to check)

$$\begin{array}{ccccc} H_n(X, A) & \xrightarrow{h} & H_n(X, U) & \xleftarrow{p} & H_n(X \setminus A, U \setminus A) \\ \downarrow q_* & & \downarrow q_* & & \downarrow q_* \\ H_n(X/A, A/A) & \xrightarrow{r} & H_n(X/A, U/A) & \xleftarrow{k} & H_n((X/A) \setminus (A/A), (U/A) \setminus (A/A)) \end{array} \quad (24.5.13)$$

Here h is an isomorphism since in the long exact sequence of the triple (X, U, A) (Proposition 24.5.5) the groups $H_n(U, A)$ are zero for all n because the deformation retraction gives a homotopy equivalence of pairs $(U, A) \simeq (A, A)$ and $H_n(A, A) = 0$. The same retraction induces a deformation retraction of U/A onto A/A , so the same argument shows that r is an isomorphism as well. The maps p, k are isomorphisms directly by excision. The rightmost map q_* is an isomorphism since q restricts to a homeomorphism on $X \setminus A$. From the commutativity of the diagram, the leftmost q_* is also an isomorphism. ■

Theorem 24.5.12. *If (X, A) is a good pair, then there is an exact sequence*

$$\cdots \tilde{H}_n(A) \xrightarrow{i_*} \tilde{H}_n(X) \xrightarrow{\pi_*} \tilde{H}_n(X/A) \xrightarrow{\partial} \tilde{H}_{n-1}(A) \rightarrow \tilde{H}_{n-1}(X) \rightarrow \cdots \rightarrow \tilde{H}_0(X/A) \rightarrow 0, \quad (24.5.14)$$

where $i : A \hookrightarrow X$ is the inclusion and $\pi : X \rightarrow X/A$ is the quotient map.

Proof. Simply combine the exact sequence of a pair (Proposition 24.5.2) with the preceding proposition. ■

Corollary 24.5.13 (Suspension Theorem). *If ΣX is the suspension of X , then $\tilde{H}_n(\Sigma X) \cong \tilde{H}_{n+1}(\Sigma X)$.*

Proof. Let $\pi : X \times [-1, 1] \rightarrow \Sigma X$ be the quotient map defining ΣX . Let $\Sigma_+ X = \pi(X \times [-\frac{1}{4}, 1])$, $\Sigma_- X = \pi(X \times [-1, \frac{1}{4}])$, $S = \pi(X \times \{-1\})$, and $N = \pi(X \times \{1\})$. Then we have the following chain of equalities

1. $\tilde{H}_i(\Sigma X) \cong H_i(\Sigma X, S)$.
2. $H_i(\Sigma X, S) \cong H_0(\Sigma X, \Sigma_- X)$ because $\Sigma_- X$ deformation retracts onto S (or alternatively from the exact sequence of the triple $(\Sigma X, \Sigma_- X, S)$ combined with $H_i(\Sigma_- X, S) = 0$).
3. $H_i(\Sigma X, \Sigma_- X) \cong H_i(\Sigma_+ X, X)$ by excising $\text{int}(\Sigma_- X)$ and by homotopy invariance.
4. $H_i(\Sigma_+ X, X) \cong \tilde{H}_{i-1}(X)$ by the long exact sequence of reduced homology for the pair $(\Sigma_+ X, X)$ combined with contractibility of $\Sigma_+ X$.

Combining all four isomorphisms we get the asserted claim. ■

Corollary 24.5.14. $\tilde{H}_m(\mathbb{S}^n) = \tilde{H}_{m-1}(\mathbb{S}^{n-1})$ and by induction the only nonzero reduced homology of a sphere is $\tilde{H}_n(\mathbb{S}^n) \cong \mathbb{Z}$.

Proof. We use induction in the dimension of the sphere. $\tilde{H}_0(\mathbb{S}^0) \cong \mathbb{Z}$ since \mathbb{S}^0 is two points. For $m > 0$ we have $\tilde{H}_m(\mathbb{S}^0) = H_i(\mathbb{S}^0) \cong H_i(*) \oplus H_i(*) = 0$. This proves the statement for $n = 0$. Now to go from n to $n + 1$, we first note that $\tilde{H}_0(\mathbb{S}^{n+1}) = 0$ because \mathbb{S}^{n+1} is connected, and for $m > 0$, $\tilde{H}_m(\mathbb{S}^{n+1}) = \tilde{H}_{m-1}(\mathbb{S}^n)$ by the Suspension Theorem. By induction, if $m = n + 1$, then this is isomorphic to \mathbb{Z} , otherwise zero. ■

Corollary 24.5.15 (Brouwer's fixed point theorem). $\partial\mathbb{D}^n$ is not a retract of \mathbb{D}^n (i.e. no continuous map $\mathbb{D}^n \rightarrow \partial\mathbb{D}^n$ that restricts to identity on the boundary). Hence every map $f : \mathbb{D}^n \rightarrow \mathbb{D}^n$ has a fixed point.

Proof. If $r : \mathbb{D}^n \rightarrow \partial\mathbb{D}^n$ is a retraction, then $r \circ i = \text{id}$ for $i : \partial\mathbb{D}^n \hookrightarrow \mathbb{D}^n$ the inclusion map. The composition $\tilde{H}_{n-1}(\partial\mathbb{D}^n) \xrightarrow{i_*} \tilde{H}_{n-1}(\mathbb{D}^n) \xrightarrow{r_*} \tilde{H}_{n-1}(\partial\mathbb{D}^n)$ is then the identity on $\tilde{H}_{n-1}(\partial\mathbb{D}^n) \cong \mathbb{Z}$. But i_* and r_* are both zero since $\tilde{H}_{n-1}(\mathbb{D}^n) = 0$, which leads to a contradiction.

An existence of a map $f : \mathbb{D}^n \rightarrow \mathbb{D}^n$ with no fixed points would let us construct a retraction by drawing the straight ray from $f(x)$ to x for all x in the ball and picking its intersection with the boundary as the image of x (this acts as the identity on the boundary). ■

Corollary 24.5.16. For a wedge sum of pointed spaces $\vee_\alpha X_\alpha$, the inclusions $i_\alpha : X_\alpha \hookrightarrow \vee_\alpha X_\alpha$ induce an isomorphism

$$\bigoplus_\alpha i_{\alpha*} : \bigoplus_\alpha \tilde{H}_n(X_\alpha) \rightarrow \tilde{H}_n\left(\bigvee_\alpha X_\alpha\right), \quad (24.5.15)$$

provided that the base points $x_\alpha \in X_\alpha$ are such that the pairs $(X_\alpha, \{x_\alpha\})$ are good.

Proof. Reduced homology is the same as homology relative to a base point, so the result follows from the long exact sequence for $(X, A) = (\bigsqcup_\alpha X_\alpha, \bigsqcup_\alpha \{x_\alpha\})$. ■

Corollary 24.5.17 (Invariance of dimension). If nonempty open sets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are homeomorphic, then $m = n$.

Proof. For $x \in U$, we have $H_k(U, U \setminus \{x\}) \cong H_k(\mathbb{R}^m, \mathbb{R}^m \setminus \{x\})$ by excision. From the long exact sequence for the pair $(\mathbb{R}^m, \mathbb{R}^m \setminus \{x\})$ we get $H_k(\mathbb{R}^m, \mathbb{R}^m \setminus \{x\}) \cong \tilde{H}_{k-1}(\mathbb{R}^m \setminus \{x\})$. Since $\mathbb{R}^m \setminus \{x\}$ deformation retracts onto the sphere \mathbb{S}^{m-1} , we have $H_k(U, U \setminus \{x\})$ is \mathbb{Z} for $k = m$ and zero otherwise. By the same reasoning $H_k(V, V \setminus \{y\})$ is \mathbb{Z} for $k = n$ and zero otherwise. A homeomorphism must induce an isomorphism between these homologies (where y is the image of x), therefore $m = n$. ■

24.6 Local homology, orientation

Definition 24.6.1 (Local homology). For a topological space X and a point $x \in X$, the local homology groups of X at x are defined to be $H_n(X, X \setminus \{x\})$, also equal by excision to $H_n(U_x, U_x \setminus \{x\})$ for any open neighborhood U_x of x .

Remark 24.6.2. Local homology is one of the ways to detect dimensions by purely topological means. For instance, we might say that X has dimension n at a point x if $H_k(X, X \setminus \{x\})$ is \mathbb{Z} for $k = n$ and zero otherwise.

Choosing a generator $\mu_x \in H_n(X, X \setminus \{x\})$ of local homology of an n -dimensional space is equivalent to choosing a local orientation. We might even construct an orientation bundle Θ whose fiber above $x \in X$ is $\Theta_x := H_n(X, X \setminus \{x\})$. The typical fiber for topological manifolds X is \mathbb{Z} and a choice of a global section $\mu \in \Gamma(\Theta)$ is equivalent to an orientation. This is the easiest way to circumvent any smoothness requirements in defining orientations.

Use Hatcher Prop 3.26-29 and to fill this out

24.7 Degree

Definition 24.7.1 (Degree on spheres). For a map $f : \mathbb{S}^n \rightarrow \mathbb{S}^n$, the induced map $f_* : \tilde{H}_n(\mathbb{S}^n) \rightarrow \tilde{H}_n(\mathbb{S}^n)$ is a homomorphism that must be of the form $f_*(\alpha) = d \cdot \alpha$ for some integer $d \in \mathbb{Z}$. This number, denoted $\deg f = d$, is called the degree of f .

Here are some properties of the degree:

1. $\deg \text{id} = 1$;
2. if f is not surjective, then $\deg f = 0$ (because $\mathbb{S}^n \setminus \{x_0\}$ is contractible and thus f is homotopic to a constant map);
3. if $f \sim g$, then $\deg f = \deg g$. In fact, the Hopf theorem states that the converse is also true;
4. $\deg f \circ g = \deg f \deg g$ since $(f \circ g)_* = f_* g_*$;
5. a reflection in any coordinate plane has degree -1 ;
6. the antipodal map ($x \mapsto -x$) on \mathbb{S}^n has degree $(-1)^{n+1}$ by virtue of being a composition of $n+1$ reflections;
7. if $f : \mathbb{S}^n \rightarrow \mathbb{S}^n$ has no fixed points, then $\deg f = (-1)^{n+1}$.

Definition 24.7.2 (Local degree). Let $f : M \rightarrow N$ be a continuous map between two connected and oriented n -dimensional topological manifolds (orientation here means a choice of a generator of the top homology group). It is said to have local degree d at point $x \in M$ (with $y = f(x)$) if the induced map in local homology $f_* : H_n(M, M \setminus \{x\}) \rightarrow H_n(N, N \setminus \{y\})$ is equivalent to multiplication by d (where $d = +1$ is the degree of an orientation-preserving local homeomorphism). We write $\deg_x f$.

Proposition 24.7.3 (Local degree formula). Suppose that for a map $f : \mathbb{S}^n \rightarrow \mathbb{S}^n$, there is a

point y such that the pre-image $f^{-1}(y)$ is finite, say $\{x_1, \dots, x_m\}$. Then

$$\deg f = \sum_{i=1}^m \deg_{x_i} f. \quad (24.7.1)$$

Proof. Exercise. ■

Corollary 24.7.4 (Fundamental theorem of algebra). *If $p(z)$ is a complex polynomial of positive degree then $p(z)$ has a zero.*

Proof. One can think of $p(z)$ as a map on the Riemann sphere $\mathbb{S}^2 \rightarrow \mathbb{S}^2$ with a fixed point at infinity, $p(\infty) = \infty$. The critical points of this map are those where $p'(z) = 0$ (because this is equivalent to the Jacobian of a holomorphic function vanishing). Clearly there are only finitely many such points because p' is a polynomial. Thus there are only finitely many points that are not regular values. The image of p is connected and p is not constant, therefore there is a regular value in the image. However, the Jacobian of a holomorphic function is always positive, so the local degree at a point mapping into a regular value must be 1. Hence the degree of p is strictly positive by the last Proposition, and hence p is not homotopic to a constant map, and its image must be all of \mathbb{S}^2 , including zero. ■

24.8 Hurewicz theorem

We have seen previously that $H_0(X)$ is free and generated by the path-connected components of X , just like $\pi_0(X)$. Computing $H_1(X)$ is a more formidable task and is the subject of the Hurewicz theorem.³⁷

We assume **WLOG** that X is path-connected and fix a base point x_0 . For any point x we denote by γ_x any path from x_0 to x (and γ_{x_0} is the constant path).

We denote by $\gamma_1 \cdot \gamma_2$ the product of paths such that $\gamma_1(1) = \gamma_2(0)$, defined as the path that traverses first γ_1 and then γ_2 at double the speed.

Lemma 24.8.1. *If γ_1 and γ_2 are paths in X such that $\gamma_1(1) = \gamma_2(0)$, then the 1-chain $\gamma_1 \cdot \gamma_2 - \gamma_1 - \gamma_2$ is a boundary.*

Proof. Define a piecewise map on the boundary of the standard 2-simplex Δ^2 by putting γ_2 on the edge $\{e_0, e_1\}$ and γ_1 on $\{e_1, e_2\}$. Then define a singular 2-simplex σ to be constant on the lines perpendicular to the edge $\{e_0, e_2\}$. This results in the path $\gamma_1 \cdot \gamma_2$ being on the edge $\{e_0, e_2\}$. Thus $\partial\sigma = \gamma_2 - \gamma_1 \cdot \gamma_2 + \gamma_1$. ■

³⁷Pronounced “Hoo-rhe-vich”

Thus in homology $[\gamma_1 + \gamma_2] = [\gamma_1 \cdot \gamma_2]$. Already we notice that when restricted to loops, the product is commutative in homology, unlike in homotopy!

Lemma 24.8.2. *If γ is a path in X , then $\gamma + \gamma^{-1}$ is a boundary. Also the constant path is a boundary.*

Proof. The constant path is the boundary of a constant 2-simplex. The rest follows from the last Lemma. ■

Lemma 24.8.3. *If γ_1 and γ_2 are paths homotopic rel ∂I (i.e. with fixed ends), then they are homologous.*

Proof. If $F : I \times I \rightarrow X$ is the homotopy, then since the edge $\{0\} \times I$ maps to a single point, F factors through the map $I \times I \rightarrow \Delta^2$ which collapses that edge to the vertex e_0 . This provides a 2-simplex σ which is γ_1 on $\{e_0, e_1\}$ and γ_2 on $\{e_0, e_2\}$ and constant on $\{e_1, e_2\}$. Then $\partial\sigma = \gamma_1 - \gamma_2 + \text{const}$. Since a constant is a boundary, so is $\gamma_1 - \gamma_2$. ■

Theorem 24.8.4 (Hurewicz). *Denote by $\tilde{\pi}_1(X, x_0)$ the abelianized fundamental group of a path-connected space X . By the last Lemma, we have a homomorphism*

$$\phi : \tilde{\pi}_1(X, x_0) \rightarrow H_1(X). \quad (24.8.1)$$

We claim that ϕ is an isomorphism.

Proof. For a path γ , put $\hat{\gamma} = \gamma_{\gamma(0)} \cdot \gamma \cdot \gamma_{\gamma(1)}^{-1}$, which is a loop based at x_0 . Define $\psi(\gamma) = [\hat{\gamma}] \in \tilde{\pi}_1(X)$. This extends to a homomorphism (because $\tilde{\pi}_1$ is abelian)

$$\psi : C_1(X) \rightarrow \tilde{\pi}_1(X). \quad (24.8.2)$$

It remains to show that ψ takes $B_1(X)$ into $0 \in \tilde{\pi}_1(X)$, so that it descends to a homomorphism ψ_* on $H_1(X)$, and that $\phi_* \circ \psi_* = 1$.

Let σ be a singular 2-simplex. Then one checks by direct computation that $\psi(\partial\sigma) = [\text{const}] = 0$ because the product of the paths making up the boundary of σ is homotopic to a point. This proves that $\psi|_{B_1(X)} = 0$.

Clearly $\psi_* \circ \phi_* = 1$ because for a loop γ , we have $\psi_* \circ \phi_*[\gamma] = [\gamma_{x_0} \cdot \gamma \cdot \gamma_{x_0}^{-1}] = [\gamma]$ by commutativity.

Finally, let σ be a 1-simplex. Then $\phi_* \circ \psi(\sigma) = [\gamma_{\sigma(0)} \cdot \sigma \cdot \gamma_{\sigma(1)-1}] = [\sigma + \gamma_{\sigma(0)} - \gamma_{\sigma(1)}]$, using the previous Lemmas. Therefore if c is a 1-chain, we have $\phi_* \circ \psi(c) = [c - \gamma_{\partial c}] \in H_1$ (where $\gamma_{\partial c}$ is the sum of the paths going from x_0 to vertices in ∂c , respecting the signs). Lastly, if c is a 1-cycle, then $\phi_* \circ \psi(c) = [c] \in H_1$, proving the theorem. ■

Similarly we can construct a Hurewicz homomorphism for higher homotopy groups:

$$h_n : \pi_n(X, x_0) \rightarrow H_n(X). \quad (24.8.3)$$

One can check that this is indeed a homomorphism, and we have the following generalization of the above theorem.

Theorem 24.8.5 (Hurewicz). *Let X be a path-connected space that is also $(n-1)$ -connected, $n \geq 2$ (that is, $\pi_k(X)$ is trivial for $k = 1, \dots, n-1$). Then $H_k(X) = 0$ for all $k < n$ and the Hurewicz map h_n is an isomorphism:*

$$h_n : \pi_n(X, x_0) \cong H_n(X). \quad (24.8.4)$$

Proof. See [19, Thm. VII.10.7]. ■

24.9 Singular vs. simplicial homology

We show now that simplicial and singular homology groups of triangulable spaces are always isomorphic, including the relative case. Let X be a triangulated space with $A \subset X$ a subcomplex, i.e. a union of simplices in X . Relative groups $H_n^\Delta(X, A)$ are defined in the same way as for singular homology, via relative chains $C_n^\Delta(X) = C_n^\Delta(X)/C_n^\Delta(A)$, and we also have the same long exact sequence of the pair (X, A) . Now we shall show that there is a canonical isomorphism $H_n^\Delta(X, A) \rightarrow H_n(X, A)$, induced by the chain map $C_n^\Delta(X, A) \rightarrow C_n(X, A)$ given by the triangulation embedding map of every simplex.

Theorem 24.9.1. *The homomorphisms $H_n^\Delta(X, A) \rightarrow H_n(X, A)$ are isomorphisms for all n and all pairs of simplicial complexes (X, A) .*

Proof. First let X be finite-dimensional and A empty. Denoting by X^k the k -skeleton of X (union of all simplices of dimension at most k), we have a commutative diagram of exact sequences:

$$\begin{array}{ccccccc} H_{n+1}^\Delta(X^k, X^{k-1}) & \xrightarrow{\quad} & H_n^\Delta(X^{k-1}) & \longrightarrow & H_n^\Delta(X^k) & \longrightarrow & H_n^\Delta(X^k, X^{k-1}) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ H_{n+1}(X^k, X^{k-1}) & \xrightarrow{\quad} & H_n(X^{k-1}) & \longrightarrow & H_n(X^k) & \longrightarrow & H_n(X^k, X^{k-1}) \\ & & & & & & \downarrow \\ & & & & & & H_{n-1}(X^{k-1}) \end{array} \quad (24.9.1)$$

Let us show that the first and the fourth vertical maps are isomorphisms for all n . The group $C_n^\Delta(X^k, X^{k-1})$ is zero for $n \neq k$ and is free abelian with basis the k -simplices of X

when $n = k$. Hence $H_n^\Delta(X^k, X^{k-1})$ is exactly the same group. The corresponding singular homology $H_n(X^k, X^{k-1})$ can be computed by considering the map

$$\Phi : \bigsqcup_\alpha (\Delta_\alpha^k, \partial\Delta_\alpha^k) \rightarrow (X^k, X^{k-1}) \quad (24.9.2)$$

formed by the embedding maps $\Delta^k \rightarrow X$ for all the k -simplices Δ_α^k in X . Since Φ induced a homeomorphism of the quotient spaces $\bigsqcup_\alpha \Delta_\alpha^k / \bigsqcup_\alpha \partial\Delta_\alpha^k \cong X^k / X^{k-1}$, it induces isomorphisms on all singular homology groups. Thus $H_n(X^k, X^{k-1})$ is zero for $k \neq n$ and for $n = k$ it is free abelian with basis represented by the relative cycles given by the embedding maps of all the k -simplices of X , in view of the fact that $H_k(\Delta^k, \partial\Delta^k)$ is generated by the identity map $\Delta^k \rightarrow \Delta^k$ (see Corollary 24.5.14).

Thus the map $H_{n+1}^\Delta(X^k, X^{k-1}) \rightarrow H_{n+1}(X^k, X^{k-1})$ is an isomorphism. By induction of this entire proof in k , we can also assume that the second and the fifth vertical maps are isomorphisms. Then by the 5-lemma 22.3.3 the middle map is an isomorphism, proving the induction step and the whole theorem for finite-dimensional X and $A = \emptyset$.

This proof is easily extended to the infinite-dimensional case due to the fact that a compact subset of X can intersect only finitely many of its simplices. This is used to show that the map $H_n^\Delta(X) \rightarrow H_n(X)$ is surjective. Namely one argues that every singular n -cycle is by definition a linear combination of finitely many singular simplices with compact images, which meet finitely many open simplices of X , and is thus contained in X^k for some finite k , where the isomorphism $H_n^\Delta(X^k) \rightarrow H_n(X^k)$ is already established. Thus the singular cycle is homologous (in X^k and thus in X) to a simplicial cycle. Injectivity is similar: if a simplicial n -cycle z is the boundary of a singular chain (i.e. its image under the map in question vanishes), then this singular chain has compact image and thus lies in some X^k , in which case by injectivity of the map $H_n^\Delta(X^k) \rightarrow H_n(X^k)$, z must be a simplicial boundary in X^k and this in X , so $[z] = 0$.

The case $A \neq \emptyset$ follows by combining the result for the absolute homology of X with an application of the 5-lemma to the analogous segment of two long exact sequences of pairs (X, A) . ■

Corollary 24.9.2. *If a space X can be triangulated by a finite simplicial complex, then all its singular homology groups are finitely generated.*

This Corollary is analogous to the theorem about the finite-dimensionality of de Rham cohomology for manifolds of finite type (with finite good cover). It also means that we have already effectively computed the singular homology groups of examples like the torus, the Möbius band, and the projective plane, by computing the simplicial homologies of their triangulations. This illustrates the value of having multiple types of equivalent homology

theories on the same space: singular homology is very general and provides powerful theoretical results but is very hard to compute directly, whereas simplicial homology can be very easy to compute in specific applications.

24.10 Axioms for homology

Based on the properties of singular homology, one can formulate a reasonable list of “axioms” for a more general homology theory of topological spaces. These axioms are known as Eilenberg–Steenrod axioms.

Definition 24.10.1 (Homology theory). A homology theory for pairs of spaces consists of a family of covariant functors $h_n : \text{TopPair} \rightarrow R\text{-Mod}$, $n \in \mathbb{Z}$, and a family of natural transformations $\partial_n : h_n \rightarrow h_{n-1} \circ \kappa$, where κ is the functor acting on objects as $\kappa(X, A) = (A, \emptyset)$ and on maps $f : (X, A) \rightarrow (Y, B)$ as $\kappa(f) = f|_A^B : (A, \emptyset) \rightarrow (B, \emptyset)$. We also write $h_n(X)$ for $h_n(X, \emptyset)$. These data must satisfy the Eilenberg–Steenrod axioms listed below.

The first three axioms are always required:

1. (Homotopy invariance.) For each homotopy F_t in TopPair we have $h_n(F_0) = h_n(F_1)$.
2. (Exact sequence.) For each pair (X, A) there is an exact sequence

$$\cdots \rightarrow h_{n+1}(X, A) \xrightarrow{\partial} h_n(A, \emptyset) \rightarrow h_n(X, \emptyset) \rightarrow h_n(X, A) \xrightarrow{\partial} \cdots \quad (24.10.1)$$

3. (Excision.) Let (X, A) be a pair and $B \subset A$ such that $\bar{B} \subset \text{int } A$. Then the inclusion $(X \setminus U, A \setminus U) \rightarrow (X, A)$ must induce an excision isomorphism $h_n(X \setminus U, A \setminus U) \cong h_n(X, A)$.

There are further optional axioms that may or may not be satisfied by specific homology theories:

4. (Dimension axiom.) $h_n(\text{point}) = 0$ for $n \neq 0$. Such a theory is called *ordinary* or *classical*, and otherwise the theory is called *extraordinary* (e.g. K-theory and cobordisms). The theory is also called *reduced* if $h_n(\text{point}) = 0$ for all n .
5. (Additivity.) h_n of a topological sum (disjoint union) of pairs is isomorphic to the direct sum of h_n ’s of the pairs.

Singular homology has two further properties that one may require:

6. (Weak equivalence.) A weak equivalence $f : X \rightarrow Y$ (i.e. a map that induces isomorphisms on homotopy groups $\pi_n(X) \cong \pi_n(Y)$) induces isomorphisms $f_* : h_\bullet(X) \cong h_\bullet(Y)$.
7. (Compact support.) For each $c \in h_n(X, A)$ there exists a map $f : (K, L) \rightarrow (X, A)$ from a pair (K, L) of compact Hausdorff spaces and $z \in h_n(K, L)$ with $f_*(z) = c$.

The last axiom holds trivially for singular homology because singular simplices are exactly such maps from the compact standard simplices into X .

Theorem 24.10.2 (Eilenberg-Steenrod). *If two sequences of covariant functors $H_n, G_n : \text{TopPair} \rightarrow \text{Ab}$, $n \geq 0$ and natural maps ∂_n, ∂'_n satisfy the Eilenberg-Steenrod axioms 1–4, then $G_n \cong H_n$ for all $n \geq 0$ (are naturally isomorphic).*

This theorem implies, for example, that the singular and simplicial homologies of a simplicial complex coincide. It should be noted that some important “homology theories” do not satisfy even the first three axioms. “Čech homology” doesn’t satisfy exactness, and “bordism theories” don’t satisfy the dimension axiom.

24.11 Cellular homology

Let us now study singular homology of CW-complexes.

Proposition 24.11.1 ([38, Lem. 39.1]). *If the CW-complex X is obtained from a subcomplex A by attaching one n -cell along $\chi : \partial\mathbb{D}^n \rightarrow A$, so $X = A \cup_{\chi} \mathbb{D}^n$, then the characteristic map $f : \mathbb{D}^n \rightarrow X$ of this cell induces isomorphisms in relative homology: $H_k(X, A) \cong H_k(\mathbb{D}^n, \partial\mathbb{D}^n)$, i.e. \mathbb{Z} when $k = n$ and 0 otherwise. Moreover, the attaching map applied to any generator of $H_n(\mathbb{D}^n, \partial\mathbb{D}^n)$ gives a generator of $H_n(X, A)$.*

Proof. Denote the n -cell $e^n = f(\mathbb{D}^n)$ and let $N \subset e^n$ be a small open neighborhood of the boundary (e.g. the image of $\partial\mathbb{D}^n \times (1 - \epsilon, 1]$ under f). By gradually shrinking N back to the boundary we obtain a homotopy equivalence $(X, A) \simeq (X, A \cup N)$ and hence isomorphisms $H_k(X, A) \cong H_k(X, A \cup N)$ by homotopy invariance. But A is in the interior of $A \cup N$, so by excision the map induced by the inclusion in homology is an isomorphism:

$$H_k(X \setminus A, N \setminus A) \cong H_k(X, A \cup N). \quad (24.11.1)$$

Now $X \setminus A$ is an open ball of dimension n and $N \setminus A$ is a collar, so clearly by collapsing this collar to a sphere we obtain an additional homotopy equivalence $(X \setminus A, N \setminus A) \simeq (e^n, \partial e^n) \simeq (\mathbb{D}^n, \partial\mathbb{D}^n)$ (the latter equivalence is induced by the attaching map f), and hence isomorphisms in homology:

$$H_k(X \setminus A, N \setminus A) \cong H_k(\mathbb{D}^n, \partial\mathbb{D}^n). \quad (24.11.2)$$

Finally, since this isomorphism is induced by the attaching map, it maps generators to generators. ■

Proposition 24.11.2. *For a CW-complex X , $H_k(X^{(n)}, X^{(n-1)}) = 0$ for $k \neq n$ and is free abelian for $k = n$, generated by the n -cells.*

Proof. This follows from Lemma 24.11.1. By applying it to all n -cells at once (which is easy because open cells are disjoint), the “total” characteristic map $f : X^{(n-1)} \cup \bigsqcup_{\alpha \in A_n} \mathbb{D}^n \rightarrow X^{(n)}$ induces homology isomorphisms

$$H_k(X^{(n)}, X^{(n-1)}) \cong H_n\left(\bigsqcup_{\alpha} \mathbb{D}^n, \bigsqcup_{\alpha} \partial \mathbb{D}^n\right). \quad (24.11.3)$$

Notice that if $k \neq n$, $H_k(\mathbb{D}^n, \partial \mathbb{D}^n) = 0$, so $H_k(\bigsqcup_{\alpha} \mathbb{D}^n, \bigsqcup_{\alpha} \partial \mathbb{D}^n) = 0$ for $k \neq n$ (the chain complex of a disjoint union is the direct sum of the respective chain complexes, and recall that the homology functors commute with direct limits), thus $H_k(X^{(n)}, X^{(n-1)}) = 0$.

For the case $k = n$, we know from singular homology that $H_n(\bigsqcup_{\alpha} \mathbb{D}^n, \bigsqcup_{\alpha} \partial \mathbb{D}^n)$ is free abelian and is isomorphic to a subgroup of \mathbb{Z}^{A_n} (where A_n is the indexing set for n -cells), hence so is $H_n(X^{(n)}, X^{(n-1)})$, with exactly one basis element for each n -cell. ■

We thus have, for example, that if X has $k < \infty$ n -cells, then $H_n(X^{(n)}, X^{(n-1)}) \cong \mathbb{Z}^k$. More generally, if the indexing set of n -cells is A_n , then $H_n(X^{(n)}, X^{(n-1)}) \cong \mathbb{Z}^{A_n}$.

Proposition 24.11.3. *If X is a CW-complex, then $H_k(X^{(n)}) = 0$ for $k > n$, and the inclusions $X^{(n)} \hookrightarrow X$ induce isomorphisms $H_k(X^{(n)}) \cong H_k(X)$ for $k < n$.*

Proof. Consider the long exact sequence of the pair $(X^{(n)}, X^{(n-1)})$:

$$\cdots \rightarrow [H_{k+1}(X^{(n)}, X^{(n-1)}) \rightarrow H_k(X^{(n-1)}) \rightarrow H_k(X^{(n)}) \rightarrow H_k(X^{(n)}, X^{(n-1)})] \rightarrow \cdots \quad (24.11.4)$$

When $k > n$, we have $k+1 \neq n$ and $k \neq n$, and by the preceding Proposition the two terms at the edges vanish and the remaining two must be isomorphic: $H_k(X^{(n-1)}) = H_k(X^{(n)})$. By iteration we get $H_k(X^{(n-1)}) = H_k(X^{(0)}) = 0$.

Now do the same for the pair $(X^{(n+1)}, X^{(n)})$:

$$\cdots \rightarrow H_{k+1}(X^{(n+1)}, X^{(n)}) \rightarrow H_k(X^{(n)}) \rightarrow H_k(X^{(n+1)}) \rightarrow H_k(X^{(n+1)}, X^{(n)}) \rightarrow \cdots \quad (24.11.5)$$

Since for $k < n$ we have $k+1 \neq n+1$ and $k \neq n+1$, by the preceding Proposition the two terms at the edges vanish, thus the remaining two are isomorphic: $H_k(X^{(n)}) \cong H_k(X^{(n+1)})$. Since $X = \varprojlim X^{(n)}$ and homology commutes with direct limits, we get $H_k(X) = H_k(X^{(n)})$. ■

It remains to find out how the boundary operator acts on singular chains $C_n(X)$. Since $H_k(X^{(n)}, X^{(n-1)})$ is generated by n -cells, it is natural to try to relate the singular boundary operator ∂ and the attaching maps. However the image of the attaching map lies in $X^{(n-1)}$, which is still a complicated space, unless we excise $X^{(n-2)}$. We thus consider the following composition of operators

$$\partial_n^{\text{CW}} : H_n(X^{(n)}, X^{(n-1)}) \xrightarrow{\delta_n} H_{n-1}(X^{(n-1)}) \xrightarrow{j_{n-1}} H_{n-1}(X^{(n-1)}, X^{(n-2)}), \quad (24.11.6)$$

where ∂_n is the connecting homomorphism of the pair $(X^{(n)}, X^{(n-1)})$ (descended from the singular boundary operator, which use the same notation), and the j_{n-1} is the map induced by “pinching out” $X^{(n-2)}$ from the same long exact sequence (but from degree $n-1$):

$$H_n(X^{(n-1)}) = 0 \rightarrow H_n(X^{(n)}) \xrightarrow{j_n} H_n(X^{(n)}, X^{(n-1)}) \xrightarrow{\partial_n} H_{n-1}(X^{(n-1)}). \quad (24.11.7)$$

Since the first term vanishes, we conclude that j_n is injective. We also have

$$\partial_n^{\text{CW}} \circ \partial_{n+1}^{\text{CW}} = j_{n-1} \circ \partial_n \circ j_n \circ \partial_{n+1} = 0, \quad (24.11.8)$$

since $\partial_n \circ j_n = 0$ is a composition of two consecutive maps in an exact sequence. We are now ready to define the cellular chain complex.

Definition 24.11.4 (Cellular homology). For a CW-complex X , its cellular chain complex is $(\mathbf{C}^{\text{CW}}, \partial^{\text{CW}})$, where $C_n^{\text{CW}} = H_n(X^{(n)}, X^{(n-1)})$ and $\partial_n^{\text{CW}} = j_{n-1} \circ \partial_n$ as introduced above. The homology groups of this complex are denoted $H_n^{\text{CW}}(X)$ and are called the cellular homology groups of X .

Theorem 24.11.5. *The cellular homology of a CW-complex X coincides with the singular homology:*

$$H_n^{\text{CW}}(X) \cong H_n(X). \quad (24.11.9)$$

Proof. First we notice that j_n and j_{n-1} are injective, therefore by exactness

$$\ker \partial_n^{\text{CW}} = \ker(j_{n-1} \circ \partial_n) = \ker \partial_n \cong \operatorname{im} j_n = H_n(X^{(n)}). \quad (24.11.10)$$

On the other hand, since j_n is injective,

$$\operatorname{im} \partial_{n+1}^{\text{CW}} = \operatorname{im}(j_n \circ \partial_{n+1}) \cong \operatorname{im} \partial_{n+1}. \quad (24.11.11)$$

Finally, the other nontrivial exact segment that we haven’t used yet comes from the long exact sequence of the pair $(X^{(n+1)}, X^{(n)})$:

$$H_{n+1}(X^{(n+1)}, X^{(n)}) \xrightarrow{\partial_{n+1}} H_n(X^{(n)}) \xrightarrow{i_n} H_n(X^{(n+1)}) \cong H_n(X) \rightarrow H_n(X^{(n+1)}, X^{(n)}) = 0, \quad (24.11.12)$$

and since the last term vanishes, i_n is surjective and the first isomorphism theorem says

$$H_n(X) \cong X_n(X^{(n)})/\ker i_m \cong H_n(X^{(n)})/\text{im } \partial_{n+1}^{\text{CW}}, \quad (24.11.13)$$

where the second equality holds by exactness.

Altogether we have

$$H_n^{\text{CW}}(X) = \ker \partial_n^{\text{CW}} / \text{im } \partial_{n+1}^{\text{CW}} \cong H_n(X^{(n)}) / \text{im } \partial_{n+1} \cong H_n(X). \quad (24.11.14)$$

■

Remark 24.11.6. We get several immediate consequences helpful in computations:

1. If X has no n -cells, then $H_n(X) = 0$. Indeed, in this case $C_n^{\text{CW}} = H_n(X^{(n)}, X^{(n-1)}) = 0$, so $H_n^{\text{CW}} = 0$.
2. If X is connected and has a single 0-cell then $\partial_1^{\text{CW}} = 0$. Indeed, since X has only one 0-cell, $C_0^{\text{CW}} \cong \mathbb{Z}$. Also, since X is connected, $H_0(X) = \mathbb{Z}$. So by the theorem above $\mathbb{Z} \cong H_0(X) = \ker \partial_0^{\text{CW}} / \text{im } \partial_1^{\text{CW}} = \mathbb{Z} / \text{im } \partial_1^{\text{CW}}$. This implies that $\text{im } \partial_1^{\text{CW}} = 0$ as claimed.
3. If X has no two cells in adjacent dimensions then $\partial_n^{\text{CW}} = 0$ for all n and $H_n(X) \cong \mathbb{Z}^{A_n}$ for all n (where A_n is the index set for n -cells). Indeed, in this case $\partial_n^{\text{CW}} = 0$ for all n , so $H_n^{\text{CW}}(X) \cong C_n^{\text{CW}} \cong \mathbb{Z}^{A_n}$.

Example 24.11.7. Recall that $\mathbb{C}P^n$ has one cell in each dimension $0, 2, 4, \dots, 2n$. So $\mathbb{C}P^n$ has no two cells in two dimensions, and we can apply the above remark:

$$H_i(\mathbb{C}P^n) = \begin{cases} \mathbb{Z}, & i = 0, 2, 4, \dots, 2n, \\ 0, & \text{otherwise.} \end{cases} \quad (24.11.15)$$

Example 24.11.8. When $n > 1$, $\mathbb{S}^n \times \mathbb{S}^n$ has one 0-cell, two n -cells, and one $2n$ -cell. Since $n > 1$, these cells are not in adjacent dimensions and we again use the above remark to compute

$$H_i(\mathbb{S}^n \times \mathbb{S}^n) = \begin{cases} \mathbb{Z}, & i = 0, 2n, \\ \mathbb{Z}^2, & i = n, \\ 0, & \text{otherwise.} \end{cases} \quad (24.11.16)$$

The rest of the section is devoted to explicitly computing the cellular boundary operator in terms of the cells. Since $\partial_n^{\text{CW}} : C_n^{\text{CW}}(X) \rightarrow C_{n-1}^{\text{CW}}(X)$ acts between two free groups generated by the sets of cells $\{e_\alpha^n\}_\alpha$ and $\{e_\beta^{n-1}\}_\beta$, we can expand it in these bases:

$$\partial_n^{\text{CW}}(e_\alpha^n) = \sum_{\beta} d_{\alpha\beta} e_\beta^{n-1}, \quad d_{\alpha\beta} \in \mathbb{Z}. \quad (24.11.17)$$

The following result provides a way of computing the coefficients.

Theorem 24.11.9. *The coefficient $d_{\alpha\beta}$ is equal to the degree of the map $\Delta_{\alpha\beta} : \mathbb{S}^{n-1} \rightarrow \mathbb{S}^{n-1}$ defined by the composition*

$$\mathbb{S}^{n-1} = \partial \mathbb{D}^n \xrightarrow{\varphi_\alpha^n} X^{(n-1)} \longrightarrow X^{(n-1)} / \left(X^{(n-2)} \sqcup \bigsqcup_{\gamma \neq \beta} e_\gamma^{(n-1)} \right) = \mathbb{S}^{n-1}, \quad (24.11.18)$$

where φ_α^n is the attaching map of e_α^n and the second arrow is the quotient map that collapses everything in $X^{(n-1)}$ except for the open cell e_β^n to a point.

Proof. We will proceed with the proof by chasing on the following diagram:

$$\begin{array}{ccccc} H_n(\mathbb{D}^n, \mathbb{S}^{n-1}) & \xrightarrow[\cong]{\partial} & \tilde{H}_{n-1}(\mathbb{S}^{n-1}) & \xrightarrow{(\Delta_{\alpha\beta})_*} & \tilde{H}_{n-1}(\mathbb{S}^{n-1}) \\ \downarrow (f_\alpha^n)_* & & \downarrow (\varphi_\alpha^n)_* & & \uparrow q_{\beta*} \\ H_n(X^{(n)}, X^{(n-1)}) & \xrightarrow{\partial_n} & \tilde{H}_{n-1}(X^{(n-1)}) & \xrightarrow{q_*} & \tilde{H}_{n-1}(X^{(n-1)} / X^{(n-2)}) \\ & \searrow \partial_n^{\text{CW}} & \downarrow j_{n-1} & & \downarrow \cong \\ & & \tilde{H}_{n-1}(X^{(n-1)} / X^{(n-2)}) & \xrightarrow{\cong} & H_n\left(\frac{X^{(n-1)}}{X^{(n-2)}}, \frac{X^{(n-2)}}{X^{(n-2)}}\right), \end{array} \quad (24.11.19)$$

where

1. f_α^n is the characteristic map (so $\partial f_\alpha^n = \varphi_\alpha^n$);
2. $q_* : \tilde{H}_{n-1}(X^{(n-1)}) \rightarrow \tilde{H}_{n-1}(X^{(n-1)} / X^{(n-2)}) = \bigoplus_\beta \tilde{H}_{n-1}(\mathbb{D}^{n-1} / \partial \mathbb{D}^{n-1})$ is induced by the quotient map $q : X^{(n-1)} \rightarrow X^{(n-1)} / X^{(n-2)}$;
3. $q_{\beta} : X^{(n-1)} / X^{(n-2)} \rightarrow \mathbb{S}^{(n-1)}$ collapses the complement of the cell e_β^{n-1} to a point, the resulting quotient sphere being identified with \mathbb{S}^{n-1} via the characteristic map f_β^{n-1} ;
4. $\Delta_{\alpha\beta} : \mathbb{S}^{n-1} \rightarrow \mathbb{S}^{n-1}$ is the composition $q_{\beta} \circ q \circ \varphi_\alpha^n$, i.e. the attaching map of e_α^n followed by the quotient map $X^{(n-1)} \rightarrow \mathbb{S}^{n-1}$ collapsing the complement of e_β^{n-1} in X^{n-1} to a point.

Note that $(\Delta_{\alpha\beta})_*$ is defined so that the top right square commutes. The upper left square is natural and therefore commutes (it is induced by the characteristic map of a cell), while the lower left triangle is part of the exact diagram defining the chain complex C_\bullet^{CW} and is defined to commute as well. The map $(f_\alpha^n)_*$ takes the generator $[\mathbb{D}^n] \in H_n(\mathbb{D}^n, \mathbb{S}^{n-1})$ to a generator of the \mathbb{Z} -summand of $H_n(X^{(n)}, X^{(n-1)})$ corresponding to e_α^n , i.e.

$$(f_\alpha^n)_*([\mathbb{D}^n]) = [e_\alpha^n]. \quad (24.11.20)$$

Since the top left square and the bottom left triangle commute, this gives that

$$\partial_n^{\text{CW}}(e_\alpha^n) = \partial_n^{\text{CW}} \circ (f_\alpha^n)_*([\mathbb{D}^n]) = j_{n-1} \circ (\varphi_\alpha^n)_* \circ \partial_n([\mathbb{D}^n]). \quad (24.11.21)$$

Looking to the bottom right square, recall that since X is a CW-complex, $(X^{(n)}, X^{(n-1)})$ is a good pair. This gives the isomorphism $H_{n-1}(X^{(n-1)}, X^{(n-2)}) \cong \tilde{H}_{n-1}(X^{(n-1)} / X^{(n-2)})$. Moreover, we also have that $\tilde{H}_{n-1}(X^{(n-1)} / X^{(n-2)}) \cong H_{n-1}(X^{(n-1)} / X^{(n-2)}, X^{(n-2)} / X^{(n-2)})$. The bottom right square commutes by definition of j_{n-1} and q_* , which combined with the commutativity of the top left square yields that

$$\partial_n^{\text{CW}}(e_\alpha^n) = q_* \circ \partial_n \circ (f_\alpha^n)_*([\mathbb{D}^n]) = q_* \circ (\varphi_\alpha^n)_* \circ \partial_n([\mathbb{D}^n]), \quad (24.11.22)$$

where formally we should precompose on the left hand side with the isomorphism $H_{n-1}(X^{(n-1)}, X^{(n-2)}) \cong \tilde{H}_{n-1}(X^{(n-1)}, X^{(n-2)})$. This last map takes the generator $[\mathbb{D}^n]$ to a linear combination of generators in $\bigoplus_\beta \tilde{H}_{n-1}(\mathbb{D}^{n-1} / \partial\mathbb{D}^{n-1})$. To see which generators it maps to, we project down to the respective β summands to obtain

$$\partial_n^{\text{CW}}(e_\alpha^n) = \sum_\beta q_{\beta*} \circ q_* \circ (\varphi_\alpha^n)_* \circ \partial_n([\mathbb{D}^n]). \quad (24.11.23)$$

As noted before, we have defined $(\Delta_{\alpha\beta})_* = q_{\beta*} \circ q_* \circ (\varphi_\alpha^n)_*$. So writing

$$\partial_n^{\text{CW}}(e_\alpha^n) = \sum_\beta (\Delta_{\alpha\beta})_* \circ \partial_n([\mathbb{D}^n]), \quad (24.11.24)$$

we see from the definition of the above maps and the fact that $\partial_n([\mathbb{D}^n])$ is a generator of $\tilde{H}_{n-1}(\mathbb{S}^{n-1})$, that $(\Delta_{\alpha\beta})_*$ is multiplication by $d_{\alpha\beta}$. ■

From now on we will drop the superscript in the name of the cellular boundary operator and just call it ∂_n whenever it doesn't collide with other notation.

Example 24.11.10. Let M_g be the closed oriented surface of genus g , with its standard CW-structure: one 0-cell, $2g$ 1-cells $\{a_1, b_1, \dots, a_g, b_g\}$, and one 2-cell attached by the product of commutators $[a_1, b_1] \cdot [a_2, b_2] \cdots [a_g, b_g]$. The associated cellular chain complex of M_g is:

$$0 \rightarrow \mathbb{Z} \xrightarrow{\partial_2} \mathbb{Z}^{2g} \xrightarrow{\partial_1} \mathbb{Z} \rightarrow 0. \quad (24.11.25)$$

Since M_g is connected and has only one 0-cell, we get that $\partial_1 = 0$. We claim that $\partial_2 = 0$ as well. Let e be the one 2-cell, so we need to compute the coefficients d_{ea_i} and d_{eb_i} in our degree formula. As the attaching map sends the generator to $a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1}$, when we collapse all 1-cells (except a_i , resp. b_i) to a point, the word defining the attaching

map reduces to $a_i a_i^{-1}$ and resp. $b_i b_i^{-1}$. Hence $d_{ea_i} = 1 - 1 = 0$. Similarly, $d_{eb_i} = 1 - 1 = 0$, for each i . Altogether,

$$\partial_2(e) = a_1 + b_1 - a_1 - b_1 + \cdots + a_g + b_g - a_g - b_g = 0. \quad (24.11.26)$$

Thus the homology groups of M_g are

$$H_n(M_g) = \begin{cases} \mathbb{Z}, & i = 0, 2, \\ \mathbb{Z}^{2g}, & i = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (24.11.27)$$

Example 24.11.11. Let N_g be the closed nonorientable surface of genus g , with its cell structure consisting of one 0-cell, g 1-cells $\{a_1, \dots, a_g\}$, and one 2-cell e attached along the word $a_1^2 \cdots a_g^2$. The cellular chain complex of N_g is given by

$$0 \rightarrow \mathbb{Z} \xrightarrow{\partial_2} \mathbb{Z}^g \xrightarrow{\partial_1} \mathbb{Z} \rightarrow 0. \quad (24.11.28)$$

As before, $\partial_1 = 0$ since N_g is connected and there is only one 0-cell. To compute $\partial_2 : \mathbb{Z} \rightarrow \mathbb{Z}^g$ we again apply the cellular boundary formula, and get

$$\partial_2(1) = (2, 2, \dots, 2) \quad (24.11.29)$$

since each a_1 appears in the attaching word with total exponent 2, which means that each map $\Delta_{\alpha\beta}$ is homotopic to the map $z \mapsto z^2$ on \mathbb{S}^1 of degree 2. In particular, ∂_2 is injective, hence $H_2(N_g) = 0$. We finally have

$$H_1(N_g) \cong \mathbb{Z}^g / \text{im } \partial_2 \cong \mathbb{Z}^g / 2\mathbb{Z} \cong \mathbb{Z}^{g-1} \oplus \mathbb{Z}_2. \quad (24.11.30)$$

Altogether,

$$H_n(N_g) = \begin{cases} \mathbb{Z}, & i = 0, \\ \mathbb{Z}^{g-1} \oplus \mathbb{Z}_2, & i = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (24.11.31)$$

Example 24.11.12. Recall that $\mathbb{R}P^n$ has a CW-structure with one cell e^k in each dimension $0 \leq k \leq n$. Moreover, the attaching map of e^k in $\mathbb{R}P^n$ is the two-fold cover projection $\varphi : \mathbb{S}^{k-1} \rightarrow \mathbb{R}P^{k-1}$. The cellular chain complex for $\mathbb{R}P^n$ is

$$0 \rightarrow \mathbb{Z} \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_1} \mathbb{Z} \rightarrow 0. \quad (24.11.32)$$

To compute the differential ∂_k , we need to compute the degree of the composite map

$$\Delta : \mathbb{S}^{k-1} \xrightarrow{\varphi} \mathbb{R}P^{k-1} \xrightarrow{q} \mathbb{R}P^{k-1}/\mathbb{R}P^{k-2} = \mathbb{S}^{k-1}. \quad (24.11.33)$$

The map Δ is a homomorphism when restricted to each component of $\mathbb{S}^{k-1} \setminus \mathbb{S}^{k-2}$, and these homeomorphisms are obtained from each other by precomposing with the antipodal map a of \mathbb{S}^{k-1} , which has degree $(-1)^k$. Hence, by our local degree formula (Proposition 24.7.3), we get that:

$$\deg \Delta = \deg \text{id} + \deg a = 1 + (-1)^k. \quad (24.11.34)$$

In particular,

$$\partial_k = \begin{cases} 0, & k = 0 \pmod{2}, \\ 2, & k = 1 \pmod{2}, \end{cases} \quad (24.11.35)$$

and therefore we obtain that

$$H_k(\mathbb{R}P^n) = \begin{cases} \mathbb{Z}_2, & k \text{ odd}, 0 < k < n, \\ \mathbb{Z}, & k = 0, \text{ or } k = n \text{ odd}, \\ 0, & \text{otherwise.} \end{cases} \quad (24.11.36)$$

Finally, note that an equivalent definition of the above map Δ is obtained by first collapsing the equatorial \mathbb{S}^{k-2} to get $\mathbb{S}^{k-1} \vee \mathbb{S}^{k-1}$, and then mapping the two copies of \mathbb{S}^{k-1} onto \mathbb{S}^{k-1} , the first one by the identity map, and the second by the antipodal map.

In summary, in real projective spaces, odd cells create new generators, even cells (except for the 0-cell) create torsion in the previous dimension.

24.12 Euler characteristic

In this section we will study CW-complexes of *finite type*, i.e. those which contain only a finite number of cells in each dimension (but they can still be infinite-dimensional). In this case each homology group is finitely generated, and from Theorem 6.2.5 we know that each finitely generated abelian group A has the form

$$A \cong \mathbb{Z}^r \oplus T(A), \quad r = \text{rank}(A), \quad (24.12.1)$$

where $T(A)$ is the torsion subgroup that can be brought to the form

$$T(A) \cong \mathbb{Z}/n_1 \oplus \mathbb{Z}/n_2 \oplus \cdots \oplus \mathbb{Z}/n_t, \quad n_1 | n_2 | \cdots | n_t, \quad (24.12.2)$$

so that the invariant factors n_1, n_2, \dots, n_t are determined uniquely. We also call them *torsion coefficients*. The following lemma follows immediately from the first isomorphism theorem.

Lemma 24.12.1. *If $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ is a short exact sequence of finitely generated abelian groups, then*

$$\text{rank } A - \text{rank } B + \text{rank } C = 0. \quad (24.12.3)$$

This is a hint that the following homological invariant is interesting.

Definition 24.12.2 (Euler characteristic). For a chain complex (\mathbf{C}, d) such that each homology group $H_n(\mathbf{C})$ is finitely generated and only finitely many of the groups $H_n(\mathbf{C})$ are nontrivial, its Euler characteristic is the natural number

$$\chi(\mathbf{C}) = \sum_{k=0}^{\infty} (-1)^k \text{rank } H_k(\mathbf{C}). \quad (24.12.4)$$

The numbers $\beta_k = \text{rank } H_k(\mathbf{C})$ are called Betti numbers. If X is a topological space then its Euler characteristic $\chi(X)$ is defined as the Euler characteristic of its singular chain complex. If X is homotopy equivalent to a simplicial or cellular complex, then the isomorphisms between singular, simplicial, and cellular homology imply that it also matches the Euler characteristic of any of the corresponding complexes.

Theorem 24.12.3. *If (\mathbf{C}, d) is a chain complex consisting of finitely generated abelian groups C_n , then*

$$\chi(\mathbf{C}) = \sum_{k=0}^{\infty} (-1)^k \text{rank } C_k. \quad (24.12.5)$$

Proof. Recall that we have two exact sequences for any complex (denoting the cycles by Z_k and the boundaries by B_k)

$$0 \rightarrow Z_k \rightarrow C_k \rightarrow B_k \rightarrow 0, \quad 0 \rightarrow B_k \rightarrow Z_k \rightarrow H_k \rightarrow 0, \quad (24.12.6)$$

so by inserting $\sum_k (\text{rank } B_k - \text{rank } B_k)$ and using the preceding lemma twice we find

$$\begin{aligned} \chi(\mathbf{C}) &= \sum_k (-1)^k (\text{rank } B_k + \text{rank } H_k + \text{rank } B_{k-1}) = \\ &= \sum_k (-1)^k (\text{rank } Z_k + \text{rank } B_{k-1}) = \sum_k (-1)^k \text{rank } C_k. \end{aligned} \quad (24.12.7)$$

■

Corollary 24.12.4 (Euler-Poincaré). *If X is homotopy equivalent to a finite CW-complex*

(or simplicial complex), then

$$\chi(X) = \sum_{k=0}^{\infty} (-1)r_k, \quad (24.12.8)$$

where r_k is the number of k -cells (k -simplices, respectively) in X .

Proof. This follows from the Theorem above combined with the homotopy invariance of singular homology and the isomorphisms between singular, simplicial, and cellular homology groups. ■

Exercise 24.7. Compute the Euler characteristic of a finite graph consisting of several connected components each of which is a tree. Now what if there are loops but the graph is planar and we triangulate the loops and include them as 2-dimensional faces?

Exercise 24.8. Compute the Euler characteristics of the simplicial complexes triangulating the circle, the sphere, and the torus. Compare the Euler characteristic of a sphere to that of a planar graph whose external “face” (the one containing the infinity in the plane) is also included.

Let us finish by briefly revisiting the Cellular Approximation Theorem 5.6.2. Given a CW-complex X of finite type, can we give a lower bound on the number of k -cells in terms of the homology of X ? As discussed above, every homology group of X can be written in the form

$$H_k(X) = \mathbb{Z}^{\beta_k} \oplus \bigoplus_{i=1}^{\tau_k} \mathbb{Z}/n_{i,k}, \quad n_{1,k}|n_{2,k}| \cdots |n_{\tau_k,k}. \quad (24.12.9)$$

Here β_i are the Betti numbers of X and τ_k are called the torsion numbers (counting the number of generators of the torsion subgroup). The minimal chain complex with homology of rank β_k in dimension k and zero elsewhere is just the complex that has \mathbb{Z}^{β_k} in degree k and zero everywhere else. The minimal chain complex of free abelian groups with homology \mathbb{Z}/n in degree k and zero elsewhere is the chain complex that vanishes outside of degrees k and $k+1$ with $\mathbb{Z} \xrightarrow{n} \mathbb{Z}$ in those degrees. Such complexes are called *elementary chain complexes*. This implies that a lower bound on the number of k -cells is

$$\beta_k + \tau_k + \tau_{k-1}, \quad (24.12.10)$$

where the first two terms comprise the generators of H_k and the last sets the relations for H_{k-1} .

These elementary complexes can be realized as the reduced cellular chains of CW-complexes (at least if $k > 0$). A bouquet of β_k copies of \mathbb{S}^k has a CW structure with one 0-cell and r k -cells, so it has homology \mathbb{Z}^{β_k} in exactly one degree. To construct a complex with cellular

chain complex given by $\mathbb{Z} \xrightarrow{\cdot^n} \mathbb{Z}$, start with \mathbb{S}^k as the k -skeleton and attach a $(k+1)$ -cell by a map of degree n . For example, when $k=1$ and $n=2$, we get $\mathbb{R}P^2$. These CW-complexes are called “Moore spaces” $M(G, k)$, where in our case $G = \mathbb{Z}$ or \mathbb{Z}/n .

The maximally efficient construction of a CW-complex for a space that is homotopy equivalent to a CW-complex of finite type can in fact be achieved, at least in the simply-connected case:

Theorem 24.12.5 (C.T.C. Wall). *Let X be a simply connected CW-complex of finite type. Then there exists a CW-complex Y with $\beta_k + \tau_k + \tau_{k-1}$ k -cells, for all k , and a homotopy equivalence $Y \rightarrow X$.*

Proposition 24.12.6. *For any sequence of abelian groups $A_k, k \geq 0$ with $A_0 = 0$, there exists a CW-complex with $\tilde{H}_k(X) = A_k$ for all $k \geq 0$.*

Proof. Let A be any abelian group. Pick generators for A . They determine a surjection from a free abelian group F_0 . The kernel F_1 of that surjection is free, being a subgroup of a free group. Write G_0 for a minimal set of generators of F_0 , and G_1 for a minimal set of generators of F_1 .

Let $k \geq 1$. Define X_k to be the wedge sum of $|G_0|$ copies of \mathbb{S}^k , so $H_k(X_k) = \mathbb{Z}[G_0]$. Now define an attaching map

$$\varphi : \bigsqcup_{b \in G_1} \mathbb{S}^k \rightarrow X_k \quad (24.12.11)$$

by specifying it on each summand \mathbb{S}^k . The generator $b \in G_1$ is given by a linear combination of generators of F_0 , say $b = \sum_{i=1}^s n_i a_i$. We want to mimic this in topology. To do this, first map $\mathbb{S}^k \rightarrow \wedge^s \mathbb{S}^k$ by pinching $(s-1)$ tangent circles to points. In homology, this map takes a generator of $H_k(\mathbb{S}^k)$ to the sum of the generators of the k -th homology of the various spheres in the bouquet. Map the i -th sphere in the bouquet to the one corresponding to a_i in X_k by a map of degree n_i . The map on the summand \mathbb{S}^k corresponding to b then is the composite of these two maps

$$\mathbb{S}^k \rightarrow \bigwedge^s \mathbb{S}^k \rightarrow \bigwedge_{a_i \in G_0} = X_k. \quad (24.12.12)$$

Altogether we get a map φ that realizes $F_1 \rightarrow F_0$ in H_k . So using it as an attaching map produces a CW-complex X with $\tilde{H}_k(X) = A$ and 0 in other degrees. This space is called a Moore space of type (A, k) , denoted $M(A, k)$.

Finally, given a full sequence of groups A_k , we only need to take a wedge sum of all spaces $M(A_k, k)$ over k . ■

Remark 24.12.7. The notation for Moore spaces is misleading because $M(G, k)$ cannot be extended to a functor $\mathbf{Ab} \rightarrow \mathbf{hTop}$ (the latter being the category of topological spaces with

| homotopy classes of continuous maps for morphisms).

25 Homological Algebra III: Derived Functors

In this section we develop more advanced tools of homological algebra necessary for dealing with cohomology with coefficients.

25.1 Injective and projective resolutions

Recall that an R -module P is called projective if the functor $\text{hom}_R(P, \cdot)$ is right-exact (it is always left-exact), i.e. if for every surjective map $A \rightarrow B$, any map $P \rightarrow B$ can be (not uniquely) factored through (lifted to) a map $P \rightarrow A$. This is a generalization of the concept of a free R -module (for free modules the lifting is unique by definition).

Proposition 25.1.1. *A left R -module P is projective iff every short exact sequence $0 \rightarrow A \rightarrow B \rightarrow P \rightarrow 0$ splits, i.e. $P \cong B/A$ implies $B \cong A \oplus P$.*

Theorem 25.1.2. *1. A left R -module is projective iff it is a direct summand of a free left R -module.
2. A finitely generated R -module is projective iff it is a direct summand of R^n for some n .
3. The direct sum of any number of projective R -modules is projective.*

For R a field or a principal ideal domain (say, \mathbb{Z}), every projective module is free (in the former case all modules are free, and in the latter all submodules of free modules are free).

Similarly, an R -module E is called injective if the functor $\text{hom}_R(\cdot, E)$ is right exact (it is always left exact), i.e. if every morphism from a submodule $A < B$ into E can always be extended (not uniquely) to a morphism $B \rightarrow E$. This is a generalization of the properties of \mathbb{Q} as a \mathbb{Z} -module (abelian group).

Proposition 25.1.3. *If a left R -module E is injective, then every short exact sequence $0 \rightarrow E \rightarrow B \rightarrow C \rightarrow 0$ splits, i.e. if E is a submodule of a module, then it is also a direct summand in it.*

Proposition 25.1.4. *1. The product of any number of injective modules is injective.
2. Every direct summand of an injective module is injective.*

Theorem 25.1.5 (Baer criterion). *An R -module E is injective iff every morphism $I \rightarrow E$, where I is an ideal in R , can be extended to a morphism $R \rightarrow E$.*

Definition 25.1.6 (Enough projectives/injectives). We say that an abelian category C has enough projectives if every object is a quotient-object of a projective one. Similarly, it has enough injectives if every object is a subobject of an injective one.

Proposition 25.1.7. *R-Mod has enough projectives and enough injectives.*

Proof. Indeed, every module M is a factor of the free module generated by all elements of M . This proves that there are enough projectives.

It requires more work to show that there are enough injectives, e.g. every module has an “injective hull” (see [27, Thm. 3.38]). ■

Definition 25.1.8 (Projective resolution). A projective resolution of an object A is an exact sequence

$$\mathbf{P} = \cdots \rightarrow P_2 \rightarrow P_1 \rightarrow P_0 \xrightarrow{\epsilon} A \rightarrow 0, \quad (25.1.1)$$

where all P_i are projective. A special case is a free resolution, and a generalization is a flat resolution, where P_i are free or flat, respectively.

If \mathbf{P} is a projective resolution of A , then its *deleted resolution* is the complex

$$\mathbf{P}_A = \cdots \rightarrow P_2 \rightarrow P_1 \xrightarrow{d_1} P_0 \rightarrow 0. \quad (25.1.2)$$

Note that the operation of deleting is reversible since $A \cong \text{coker } d_1$.

Proposition 25.1.9. *Every R-module A has a projective (in fact, free) resolution.*

Proof. There is a free module F_0 with an exact sequence

$$0 \rightarrow K_1 \rightarrow F_0 \xrightarrow{\epsilon} A \rightarrow 0. \quad (25.1.3)$$

Here, K_1 is the module of *relations* on A . Similarly there is a free F_1 with an exact sequence

$$0 \rightarrow K_2 \rightarrow F_1 \xrightarrow{\epsilon_1} K_1 \rightarrow 0. \quad (25.1.4)$$

These higher K_n 's are called *syzygies* of A (relations on relations etc.). Define $d_1 = i_1 \circ \epsilon_1$. Then clearly $\text{im } d_1 = K_1 = \ker \epsilon$ and $\ker d_1 = K_2$, giving a spliced exact sequence

$$F_1 \xrightarrow{d_1} F_0 \xrightarrow{\epsilon} A \rightarrow 0. \quad (25.1.5)$$

Iterating this construction, we get a free resolution (which is also projective and flat). ■

Corollary 25.1.10. *In a category with enough projectives, every object has a projective resolution (by repeating the same proof).*

Definition 25.1.11 (Injective resolution). An injective resolution of an object A is an exact sequence

$$\mathbf{E} = 0 \rightarrow A \xrightarrow{\eta} E^0 \rightarrow E^1 \rightarrow E^2 \rightarrow \dots \quad (25.1.6)$$

in which all E^i are injective.

We also have the *deleted injective resolution*, which is the complex

$$\mathbf{E}^A = 0 \rightarrow E^0 \xrightarrow{d^0} E^1 \rightarrow E^2 \rightarrow \dots \quad (25.1.7)$$

Again, $A \cong \ker d^0$, so a deleted resolution contains the same information as the full one.

Proposition 25.1.12. *Every R -module A has an injective resolution.*

Proof. Relying on the fact that there are enough injectives in $R\text{-Mod}$ ([27, Thm. 3.38]), we can repeat a dualized version of the same proof as in the last Proposition. The analogs of syzygies here are called *cosyzygies*. ■

Corollary 25.1.13. *In a category with enough injectives, every object has an injective resolution.*

Theorem 25.1.14 (Comparison theorem for projective resolutions). *Let $f : A \rightarrow A'$ be a morphism in a category with enough projectives and consider a diagram*

$$\begin{array}{ccccccc} \cdots & \longrightarrow & P_1 & \xrightarrow{d_1} & P_0 & \xrightarrow{\epsilon} & A \longrightarrow 0 \\ & & & & & & \downarrow f \\ \cdots & \longrightarrow & X_1 & \xrightarrow{d'_1} & X_0 & \xrightarrow{\epsilon'} & A' \longrightarrow 0 \end{array} \quad (25.1.8)$$

where the top row is a complex with all P_i projective, and the bottom row is exact. Then there exists a chain map between the rows extending f , and it is unique up to homotopy:

$$\begin{array}{ccccccc} \cdots & \longrightarrow & P_1 & \longrightarrow & P_0 & \longrightarrow & A \longrightarrow 0 \\ & & \downarrow f_1 & & \downarrow f_0 & & \downarrow f \\ \cdots & \longrightarrow & X_1 & \longrightarrow & X_0 & \longrightarrow & A' \longrightarrow 0 \end{array} \quad (25.1.9)$$

Proof. We prove the existence of f_n by induction in n . For the base step consider the diagram

$$\begin{array}{ccccc} P_0 & & & & \\ f_0 \downarrow & & f \circ \epsilon & & \\ X'_0 & \xrightarrow{\epsilon'} & A' & \longrightarrow & 0 \end{array} \quad (25.1.10)$$

Since ϵ' is surjective and P_0 is projective, there is a f_0 .

For the inductive step, we first notice that $\text{im}(f_n \circ d_{n+1}) \subset \text{im } d'_{n+1}$ (by exactness of the bottom row this follows from $d'_n f_n d_{n+1} = f_{n-1} d_n d_{n+1} = 0$), so that we have the diagram

$$\begin{array}{ccccc} P_{n+1} & & & & \\ f_{n+1} \downarrow & & f_n \circ d_{n+1} & & \\ X_{n+1} & \xrightarrow{d'_{n+1}} & \text{im } d'_{n+1} & \longrightarrow & 0 \end{array} \quad (25.1.11)$$

and then projectivity of P_{n+1} guarantees a map f_{n+1} .

Finally, given another chain map $h : \mathbf{P}_A \rightarrow \mathbf{P}'_{A'}$ with $\epsilon' \circ h_0 = f \circ \epsilon$, one can construct maps $s_n : P_n \rightarrow P'_{n+1}$ such that

$$h_n - f_n = d'_{n+1} s_n + s_{n-1} d_n. \quad (25.1.12)$$

See [27, Thm. 6.16] for the details of this construction. ■

Corollary 25.1.15. *Any two projective resolutions of an object are homotopy equivalent.*

Proof. Since now \mathbf{X} is also a projective resolution, and $A = A'$ so that $f = \text{id}$ is an isomorphism, we have chain maps f_\bullet and g_\bullet in both directions. Taking compositions, we get chain maps $fg : \mathbf{X} \rightarrow \mathbf{X}$ and $gf : \mathbf{P} \rightarrow \mathbf{P}$. However, by the last theorem these are unique up to homotopy, and clearly the identity chain maps also work. Thus both fg and gf are homotopic to identities. ■

Corollary 25.1.16 (Comparison theorem for injective resolutions). *A dual statement holds for injective resolutions (then the top row is exact and the bottom row is a complex of injectives).*

Definition 25.1.17 (Chain map over a morphism). We say that the chain map between two deleted resolutions $f_\bullet : \mathbf{P}_A \rightarrow \mathbf{P}'_{A'}$ is *over* the morphism $f : A \rightarrow A'$ if $f \circ \epsilon = \epsilon' \circ f_0$. We have established that for every f and two given deleted resolutions there is a chain map over f , and it is unique up to homotopy.

25.2 Left derived covariant functors

As we already know, short exact sequences under the action of a particular functor do not, in general, remain exact. Many relevant functors (say, covariant for now) preserve exactness only on the left (“left-exact functor”) or the right (“right-exact functor”) sides of the sequences. We are now ready to define the objects that measure the non-exactness of a given functor (e.g. a hom-functor or a tensor product functor).

Definition 25.2.1 (Left derived functors). Let $F : \mathcal{C} \rightarrow \mathcal{D}$ be an additive covariant functor between abelian categories and let \mathcal{C} have enough projectives. For an object $A \in \text{Ob}(\mathcal{C})$ choose an arbitrary deleted projective resolution \mathbf{P}_A :

$$\cdots \rightarrow P_2 \rightarrow P_1 \rightarrow P_0 \rightarrow 0. \quad (25.2.1)$$

Then $F(\mathbf{P}_A)$ is a complex in \mathcal{D} . We define the left derived functors $L_n F$ by

$$(L_n F)(A) = H_n(F(\mathbf{P}_A)), \quad n \geq 0. \quad (25.2.2)$$

Their (covariant) actions on morphisms $f : A \rightarrow B$ are defined using the chain maps from the comparison theorem,

$$(L_n F)(f) = H_n(F(f_n)) = F(f)_{*n}. \quad (25.2.3)$$

Proposition 25.2.2. *$L_n F$ is indeed a functor for each $n \geq 0$, which is moreover additive and covariant. In addition, alternative choices of resolutions of all objects lead to different but naturally isomorphic functors $L_n F$.*

Proof. The first statement is the functoriality of $L_n F$ when a resolution is fixed for each object in the category. For this, we recall that for every morphism $f : A \rightarrow A'$, the chain map over it is unique up to homotopy, therefore $H_n(Ff)$ is independent of the choice of this chain map. Additivity is a simple exercise.

For the second statement, we need to show that a different choice of resolutions for all objects of the category produces another functor $\tilde{L}_n F$ that is naturally isomorphic to the original $L_n F$. The crucial step is the uniqueness of the projective resolution up to homotopy equivalence which guarantees that the homology groups for different choices of resolutions are isomorphic. The rest is checking that the squares formed by these isomorphisms commute, see also [27, Prop. 6.20]. ■

The trick in practice therefore is finding the resolution that simplifies the computations the most.

Lemma 25.2.3. *If the additive covariant functor F is right exact, then $L_0F \cong F$ (naturally isomorphic functors).*

Proof. Let \mathbf{P} be a projective resolution of A which under the action of F maps to a right exact sequence $\cdots \rightarrow F(P_1) \xrightarrow{d_1} F(P_0) \rightarrow F(A) \rightarrow 0$. Therefore $F(A) \cong F(P_0)/\text{im } F(d_1)$. On the other hand, $(L_0F)(A) = H_0(F(\mathbf{P}_A)) = \ker F(d_0)/\text{im } F(d_1) = F(P_0)/\text{im } F(d_1) \cong F(A)$. Therefore the objects produced by F and L_0F are isomorphic. It is trivial to check that squares produced by these isomorphisms commute, which means that the two functors are naturally isomorphic. ■

We already know that an exact sequence of complexes gives rise to an exact sequence in homology. However, what we really want is an exact sequence of objects in the original category. Therefore we need to find a way to produce a long exact sequence.

Lemma 25.2.4 (Horseshoe lemma). *Given a diagram in an abelian category with enough projectives*

$$\begin{array}{ccccccc} P'_1 & & & P''_1 & & & \\ \downarrow & & & \downarrow & & & \\ P'_0 & & & P''_0 & & & \\ \downarrow \epsilon' & & & \downarrow \epsilon'' & & & \\ 0 \longrightarrow A' \xrightarrow{i} A \xrightarrow{q} A'' \longrightarrow 0, & & & & & & \end{array} \quad (25.2.4)$$

where the columns are projective resolutions, if the bottom row is exact, then there is a projective resolution of A and horizontal chain maps such that the three columns form an exact sequence of complexes.

The dual version with injective resolutions holds as well.

Proof. Define $Q_0 = P'_0 \oplus P''_0$. It is projective because P'_0 and P''_0 are. Define $i_0 : P'_0 \rightarrow Q_0$ by $x \mapsto (x', 0)$ and $q_0 : Q_0 \rightarrow P''_0$ by $(x', x'') \mapsto x''$. Clearly the sequence

$$0 \rightarrow P'_0 \xrightarrow{i_0} Q_0 \xrightarrow{q_0} P''_0 \rightarrow 0 \quad (25.2.5)$$

is exact. Since P''_0 is projective, there is a map $\sigma : P''_0 \rightarrow A$ with $q\sigma = \epsilon''$. Now define $\epsilon : Q_0 \rightarrow A$ by $(x', x'') \mapsto i\epsilon'(x') + \sigma(x'')$. This ensures that the square

$$\begin{array}{ccc} Q_0 & \xrightarrow{q_0} & P'_0 \\ \downarrow \epsilon & \swarrow \sigma & \downarrow \epsilon'' \\ A & \xrightarrow{q} & A'' \end{array} \quad (25.2.6)$$

commutes. The surjectivity of ϵ follows from the 5-lemma. Now we define $V_0 = \ker \epsilon$, $K'_0 = \ker \epsilon'$ and $K''_0 = \ker \epsilon''$, and establish the existence of maps $K'_0 \rightarrow V_0 \rightarrow K''_0$ such that the resulting 3×3 diagram commutes:

$$\begin{array}{ccccccc}
& 0 & & 0 & & 0 & \\
& \downarrow & & \downarrow & & \downarrow & \\
0 & \longrightarrow & K'_0 & \longrightarrow & V_0 & \longrightarrow & K''_0 \longrightarrow 0 \\
& \downarrow & & \downarrow & & \downarrow & \\
0 & \longrightarrow & P'_0 & \xrightarrow{i_0} & Q_0 & \xrightarrow{q_0} & P''_0 \longrightarrow 0 \\
& & \downarrow \epsilon' & & \downarrow \epsilon & & \downarrow \epsilon'' \\
0 & \longrightarrow & A' & \xrightarrow{i} & A & \xrightarrow{q} & A'' \longrightarrow 0 \\
& & \downarrow & & \downarrow & & \downarrow \\
& 0 & & 0 & & 0 &
\end{array} \tag{25.2.7}$$

Furthermore, by the 3×3 -lemma 22.3.9, the top row is exact. Thus we have constructed the first row $0 \rightarrow P'_0 \rightarrow Q_0 \rightarrow P''_0 \rightarrow 0$ of the desired diagram.

Now we extend this construction upwards by induction. Assuming the bottom n rows have been constructed (with vertical maps d'_n, δ_n and d''_n respectively), let $V_n = \ker(Q_n \rightarrow Q_{n-1})$, $K'_n = \ker d'_n$ and $K''_n = \ker d''_n$. As in the base step, we have a commuting diagram with exact rows and columns

$$\begin{array}{ccccccc}
& 0 & & 0 & & 0 & \\
& \downarrow & & \downarrow & & \downarrow & \\
0 & \longrightarrow & K'_{n+1} & \longrightarrow & V_{n+1} & \longrightarrow & K''_{n+1} \longrightarrow 0 \\
& \downarrow & & \downarrow & & \downarrow & \\
0 & \longrightarrow & P'_{n+1} & \xrightarrow{i_{n+1}} & Q_{n+1} & \xrightarrow{q_{n+1}} & P''_{n+1} \longrightarrow 0 \\
& & \downarrow d'_n & & \downarrow \delta_{n+1} & & \downarrow d''_{n+1} \\
0 & \longrightarrow & K'_n & \longrightarrow & V_n & \longrightarrow & K''_n \longrightarrow 0 \\
& & \downarrow & & \downarrow & & \downarrow \\
& 0 & & 0 & & 0 &
\end{array} \tag{25.2.8}$$

Defining $\delta_{n+1} : Q_{n+1} \rightarrow Q_n$ as the composition $Q_{n+1} \rightarrow V_n \rightarrow Q_n$, we splice this diagram into the existing one to produce the next row. ■

Remark 25.2.5. Note that while every row of the resulting diagram splits by projectivity of P_n'' , the splitting maps $P_n'' \rightarrow P_n$ don't form a chain map, i.e. the final sequence of complexes itself does not necessarily split.

The next theorem finally explains in what sense derived functors of F measure the non-exactness of F . Namely, their values is what needs to be put on the left of the image of a short exact sequence under F to make it exact again.

Theorem 25.2.6 (Long exact sequence of left derived functors). *Given a covariant additive functor $F : C \rightarrow \mathcal{D}$ as above, for every short exact sequence in C*

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0 \quad (25.2.9)$$

there is a long exact sequence of left derived functors

$$\cdots \rightarrow (L_1 F)(C) \xrightarrow{\delta_1} (L_0 F)(A) \rightarrow (L_0 F)(B) \rightarrow (L_0 F)(C) \rightarrow 0. \quad (25.2.10)$$

Proof. Let \mathbf{P}' and \mathbf{P}'' be the chosen projective resolutions of A and C (the ones used in defining $L_n F$). By the Horseshoe Lemma, there is a projective resolution $\tilde{\mathbf{P}}$ of B (not necessarily coinciding with the chosen resolution, hence the tilde) giving an exact sequence of complexes

$$0 \rightarrow \mathbf{P}'_A \xrightarrow{j} \tilde{\mathbf{P}}_B \xrightarrow{q} \mathbf{P}''_C \rightarrow 0. \quad (25.2.11)$$

Here j is a chain map over i and q over p . Applying F gives a sequence of complexes

$$0 \rightarrow F\mathbf{P}'_A \xrightarrow{Fj} F\tilde{\mathbf{P}}_B \xrightarrow{Fq} F\mathbf{P}''_C \rightarrow 0. \quad (25.2.12)$$

Since each row $0 \rightarrow P'_n \rightarrow \tilde{P}_n \rightarrow P''_n \rightarrow 0$ is a split exact sequence (by projectivity of P''_n), and additive functors preserve split short exact sequences, the sequence above is also exact. Therefore we have a long exact sequence

$$\cdots \rightarrow H_n(F\mathbf{P}'_A) \xrightarrow{(Fj)_*} H_n(F\tilde{\mathbf{P}}_B) \xrightarrow{(Fq)_*} H_n(F\mathbf{P}''_C) \xrightarrow{\delta_n} H_{n-1}(F\mathbf{P}'_A) \rightarrow \cdots, \quad (25.2.13)$$

which translates to

$$\cdots \rightarrow (L_n F)(A) \xrightarrow{(Fj)_*} (\tilde{L}_n F)(B) \xrightarrow{(Fq)_*} (L_n F)(C) \xrightarrow{\delta_n} (L_{n-1} F)(A) \rightarrow \cdots. \quad (25.2.14)$$

This sequence indeed terminates with 0 because $L_{-1} F$ is zero by construction.

Using the homotopy invariance property of derived functors, one can replace \tilde{L}_n here by L_n without breaking the exactness of the sequence, see [27, Thm. 6.27] for details. ■

Proposition 25.2.7 (Naturality of δ). *The long exact sequence of left derived functors (in particular the connecting homomorphism δ) is natural with respect to morphisms between short exact sequences of objects.*

Proof. This is just a version of Theorem 24.2.14. ■

25.3 Left derived contravariant functors

This case will not come up in practice, so we quickly go over it for completeness.

Definition 25.3.1 (Left derived contravariant functors). Let $F : C \rightarrow \mathcal{D}$ be an additive contravariant functor between abelian categories and let C have enough injectives. For an object $A \in \text{Ob}(C)$, choose an arbitrary deleted injective resolution \mathbf{E}^A :

$$0 \rightarrow E^0 \rightarrow E^1 \rightarrow E^2 \rightarrow \cdots. \quad (25.3.1)$$

Then $F(\mathbf{E}^A)$ is a complex in \mathcal{D} . We define the left derived contravariant functors $L_n F$ by

$$(L_n F(A) = H_n(F(\mathbf{E}^A)), n \geq 0). \quad (25.3.2)$$

Their (contravariant) actions on morphisms $f : A \rightarrow B$ are defined using the chain maps from the comparison theorem,

$$(L_n F)(f) = H_n(F(f_n)) = F(f)_{*n}. \quad (25.3.3)$$

By repeating the derivations from the covariant case we get the following properties.

Theorem 25.3.2. *Given a contravariant additive functor F ,*

1. *$L_n F$ is a contravariant additive functor for each $n \geq 0$, and different choices of the injective resolution lead to naturally isomorphic functors $L_n F$.*
2. *If F is right exact, there is a natural isomorphism $L_0 F \cong F$.*
3. *A short exact sequence of objects $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ gives rise to a long exact sequence*

$$\cdots (L_1 F)(A) \xrightarrow{\delta_1} (L_0 F)(C) \rightarrow (L_0 F)(B) \rightarrow (L_0 F)(A) \rightarrow 0. \quad (25.3.4)$$

4. *The long exact sequence above is functorial with respect to morphisms between short exact sequences.*

25.4 Right derived functors

Definition 25.4.1 (Right derived covariant functors). Let $F : C \rightarrow \mathcal{D}$ be an additive covariant functor between abelian categories and let C have enough injectives. For an object $A \in \text{Ob}(C)$, choose an arbitrary deleted injective resolution \mathbf{E}^A :

$$0 \rightarrow E^0 \rightarrow E^1 \rightarrow E^2 \rightarrow \dots . \quad (25.4.1)$$

Then $F(\mathbf{E}^A)$ is a complex in \mathcal{D} . We define the right derived covariant functors $R_n F$ by

$$(R^n F)(A) = H^n(F(\mathbf{E}^A)), \quad n \geq 0. \quad (25.4.2)$$

Their (contravariant) actions on morphisms $f : A \rightarrow B$ are defined using the chain maps from the comparison theorem,

$$(R^n F)(f) = H^n(F(f_n)) = F(f)^{*n}. \quad (25.4.3)$$

Theorem 25.4.2. Given a covariant additive functor F ,

1. $R^n F$ is a covariant additive functor for each $n \geq 0$, and different choices of the injective resolution lead to naturally isomorphic functors $R^n F$.
2. If F is left exact, there is a natural isomorphism $R^0 F \cong F$.
3. A short exact sequence of objects $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ gives rise to a long exact sequence

$$0 \rightarrow (R^0 F)(A) \rightarrow (R^0 F)(B) \rightarrow (R^0 F)(C) \xrightarrow{\delta^0} (R^1 F)(A) \rightarrow \dots . \quad (25.4.4)$$

4. The long exact sequence above is functorial with respect to morphisms between short exact sequences.

Definition 25.4.3 (Right derived contravariant functors). Let $F : C \rightarrow \mathcal{D}$ be an additive covariant functor between abelian categories and let C have enough projectives. For an object $A \in \text{Ob}(C)$ choose an arbitrary deleted projective resolution \mathbf{P}_A :

$$\dots \rightarrow P_2 \rightarrow P_1 \rightarrow P_0 \rightarrow 0. \quad (25.4.5)$$

Then $F(\mathbf{P}_A)$ is a complex in \mathcal{D} . We define the right derived functors $R^n F$ by

$$(R^n F)(A) = H^n(F(\mathbf{P}_A)), \quad n \geq 0. \quad (25.4.6)$$

Their (contravariant) actions on morphisms $f : A \rightarrow B$ are defined using the chain maps from the comparison theorem,

$$(R^n F)(f) = H^n(F(f_n)) = F(f)^{*n}. \quad (25.4.7)$$

Theorem 25.4.4. *Given a contravariant additive functor F ,*

1. *$R^n F$ is a contravariant additive functor for each $n \geq 0$, and different choices of the projective resolution lead to naturally isomorphic functors $R^n F$.*
2. *If F is left exact, there is a natural isomorphism $R^0 F \cong F$.*
3. *A short exact sequence of objects $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ gives rise to a long exact sequence*

$$0 \rightarrow (R^0 F)(C) \rightarrow (R^0 F)(B) \rightarrow (R^0 F)(A) \xrightarrow{\delta^0} (R^1 F)(C) \rightarrow \dots \quad (25.4.8)$$

4. *The long exact sequence above is functorial with respect to morphisms between short exact sequences.*

25.5 Tor functors

By far the most common application of derived functors is to the functors Hom_R and \otimes_R . These are, generally speaking, functors of two arguments, and it is possible to define their derived functors with respect to both arguments at once. However, it is easier to apply the existing construction to the functors obtained by fixing one of the arguments. The natural isomorphism of these two approaches is proven e.g. in [27].

Recall that the tensor product $A \otimes_R B \in \mathbf{Ab}$ of a right R -module A and a left R -module B can be defined by the universal property that any R -bilinear map $f : A \times B \rightarrow G$ (these are properly called *R-balanced products*: they have to be additive on both arguments and also satisfy $f(a \cdot r, b) = f(a, r \cdot b)$ for all $a \in A, b \in B, r \in R$) into any abelian group G uniquely factors through $A \otimes_R B$. Alternatively, the tensor product can be defined as the adjoint of the hom-functor (here we use the fact that \mathbb{Z} -modules are the same thing as abelian groups, so we can write $\text{Hom}_{\mathbb{Z}}$ to denote the morphism sets in \mathbf{Ab}):

$$\text{Hom}_{\mathbb{Z}}(A \otimes_R B, G) \cong \text{Hom}_R(A, \text{Hom}_{\mathbb{Z}}(B, G)). \quad (25.5.1)$$

If R is commutative (and hence right and left R -modules are isomorphic), then $A \otimes_R B$ can be given a natural structure of an R -module as well by defining $r \cdot (a \otimes b) = (r \cdot a) \otimes b = a \otimes (r \cdot b)$ (and the universal property can be adjusted by replacing abelian groups G with R -modules). In particular, if $R = \mathbb{Z}$, then we get the tensor product of abelian groups $\otimes : \mathbf{Ab} \times \mathbf{Ab} \rightarrow \mathbf{Ab}$.

Example 25.5.1. 1. If G is an torsion abelian group, i.e. every element has finite order, then $\mathbb{Q} \otimes_{\mathbb{Z}} G = 0$. Indeed, every element of this group can be written as $\sum_i q_i \otimes g_i$, and assuming $n_i \cdot g_i = 0$, this is the same as $\sum_i (q_i/n_i) \otimes n_i g_i = 0$.

2. $\mathbb{Z}_m \otimes_{\mathbb{Z}} \mathbb{Z}_n \cong \mathbb{Z}_{\gcd(m,n)}$. More generally, for any pair of ideals $I, J \trianglelefteq R$, we have

$$R/I \otimes_R R/J = \frac{R/J}{I(R/J)} = \frac{R/J}{(I+J)/J} = R/(I+J). \quad (25.5.2)$$

3. Tensor products with the cyclic groups can be used to control the order of elements of groups. If G is an abelian group, then $G \otimes_{\mathbb{Z}} \mathbb{Z}_n = G/nG$. All multiples of n in this group are zero.
4. $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q} \otimes_{\mathbb{Q}} \mathbb{Q}$.

Definition 25.5.2 (Left Tor functor). Let B be a left R -module and consider the right exact functor $\underline{} \otimes_R B : \text{Mod-}R \rightarrow \text{Ab}$, and define

$$\text{Tor}_n^R(\underline{}, B) = L_n(\underline{} \otimes_R B). \quad (25.5.3)$$

Thus, if $\mathbf{P} = \cdots \rightarrow P_2 \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} A \rightarrow 0$ is a projective resolution of a right R -module A , then

$$\text{Tor}_n^R(A, B) = H_n(\mathbf{P}_A \otimes_R B) = \frac{\ker(d_n \otimes 1_B)}{\text{im}(d_{n+1} \otimes 1_B)}. \quad (25.5.4)$$

Applying the general properties of left derived covariant functors, we get the following theorem.

Theorem 25.5.3. *The left Tor functor $\text{Tor}_n^R(\underline{}, M)$ has the following properties:*

1. *There is a natural isomorphism $\text{Tor}_0^R(A, M) \cong A \otimes_R M$.*
2. *A short exact sequence $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ of right R -modules gives rise to a long exact sequence*

$$\cdots \rightarrow \text{Tor}_1^R(B, M) \rightarrow \text{Tor}_1^R(C, M) \xrightarrow{\delta_1} A \otimes_R M \rightarrow B \otimes_R M \rightarrow C \otimes_R M \rightarrow 0, \quad (25.5.5)$$

and the connecting homomorphisms δ_i are natural.

3. *If A is projective (i.e. $0 \rightarrow A \xrightarrow{1} A \rightarrow 0$ is a projective resolution), then $\text{Tor}_n^R(A, M) = 0$, $n \geq 1$.*

Theorem 25.5.4. *Fixing a left R -module M , and a sequence of covariant functors $F_n : \text{Mod-}R \rightarrow \text{Ab}$, $n \geq 0$, which satisfy the properties of Tor_n^R established in Theorem 25.5.3, there are natural isomorphisms $F_n \cong \text{Tor}_n^R(\underline{}, M)$.*

Proof. Induction in n . The base $n = 0$ holds by property 1. Now let us look at the induction

step to $n = 1$. Consider a short exact sequence

$$0 \rightarrow B \rightarrow P \rightarrow A \rightarrow 0 \quad (25.5.6)$$

with a projective P . We have two long exact sequences

$$\begin{array}{ccccccc} F_1(P) = 0 & \longrightarrow & F_1(A) & \longrightarrow & B \otimes_R M & \rightarrow & P \otimes_R M \rightarrow \cdots \\ \downarrow 0 & & \downarrow & & \downarrow 1 & & \downarrow 1 \\ \mathrm{Tor}_1^R(P, M) = 0 & \rightarrow & \mathrm{Tor}_1^R(A, M) & \rightarrow & B \otimes_R M & \rightarrow & P \otimes_R M \rightarrow \cdots \end{array} \quad (25.5.7)$$

with a commutative square on the right. Note that $F_1(A)$ is the kernel of the morphism coming out of it, and $\mathrm{Tor}_1^R(B, M)$ is also the kernel of the rightward morphism from it. By Lemma 22.3.1, there is a corresponding morphism $F_1(A) \rightarrow \mathrm{Tor}_1^R(B, M)$ that makes the left square commute. Moreover, since the two vertical morphisms that we have are isomorphisms, this new morphism is also an isomorphism by the 5-lemma 22.3.3.

Now assume we have an isomorphism in degree $n - 1$. We will apply a trick called *dimension shifting*. We consider the two corresponding segments of the above long sequences:

$$\begin{array}{ccccccc} F_n(P) = 0 & \longrightarrow & F_n(A) & \longrightarrow & F_{n-1}(B) & \longrightarrow & F_{n-1}(P) = 0 \\ \downarrow 0 & & \downarrow & & \downarrow & & \downarrow 0 \\ \mathrm{Tor}_n^R(P, M) = 0 & \rightarrow & \mathrm{Tor}_n^R(A, M) & \rightarrow & \mathrm{Tor}_{n-1}^R(B, M) & \rightarrow & \mathrm{Tor}_{n-1}^R(P, M) = 0. \end{array} \quad (25.5.8)$$

The nonzero solid vertical morphism here is an isomorphism by the induction assumption, and the morphisms $F_n(A) \rightarrow F_{n-1}(B)$ and $\mathrm{Tor}_n^R(A, M) \rightarrow \mathrm{Tor}_{n-1}^R(B, M)$ are isomorphisms as well, therefore there exists a morphism $F_n(A) \rightarrow \mathrm{Tor}_n^R(A, M)$ that makes the diagram commute and it is an isomorphism. ■

Definition 25.5.5 (Right Tor functor). Let A be a right R -module and consider the right exact functor $A \otimes_R - : R\text{-Mod} \rightarrow \mathbf{Ab}$, and define

$$\mathrm{tor}_n^R(A, -) = L_n(A \otimes_R -). \quad (25.5.9)$$

Thus, if $\mathbf{Q} = \cdots \rightarrow Q_2 \xrightarrow{d_2} Q_1 \xrightarrow{d_1} Q_0 \xrightarrow{\epsilon} B \rightarrow 0$ is a projective resolution of a right R -module B , then

$$\mathrm{tor}_n^R(A, B) = H_n(A \otimes_R \mathbf{Q}_B) = \frac{\ker(1_A \otimes d_n)}{\mathrm{im}(1_A \otimes d_{n+1})}. \quad (25.5.10)$$

Theorem 25.5.6. *If \mathbf{P}_A and \mathbf{Q}_B are deleted projective resolutions of A and B , respectively, then there are natural isomorphisms*

$$H_n(\mathbf{P}_A \otimes_R B) \cong H_n(A \otimes_R \mathbf{Q}_B) \quad (25.5.11)$$

as functors of A and B . In particular,

$$\mathrm{Tor}_n^R(A, B) \cong \mathrm{tor}_n^R(A, B). \quad (25.5.12)$$

Proof. Any proof of this theorem is quite tedious, but the most intuitive way of proving it is by considering the \mathbb{Z} -graded product of the two complexes (a *double complex*), $(A \otimes B)_n = \oplus_{p+q=n} A_p \otimes B_q$. The differential here can be defined via operators $D_{p,q} : A_p \otimes B_q \rightarrow (A \otimes B)_{n-1}$ that act by

$$X \otimes Y \mapsto d_p X \otimes Y + (-1)^p X \otimes d_q Y. \quad (25.5.13)$$

Then $D_n = \oplus_{p+q=n} D_{p,q}$. It is easy to check that $D^2 = 0$, since for $x \in A_p, y \in B_q$,

$$D^2(x \otimes y) = d^2 x \otimes y + (-1)^p dx \otimes dy + (-1)^{p-1} dx \otimes dy + x \otimes d^2 y = 0, \quad (25.5.14)$$

so we have a complex and can define homology groups. $H_n(A \otimes B)$ thus ends up being explicitly expressed in terms of the homologies of the resolutions of A and B . Then one proves separate isomorphisms between each of the two homology groups above and $H_n(\mathbf{P}_A \otimes \mathbf{Q}_B)$. We omit the details.

For a detailed proof that avoids this construction see [27, Thm. 6.32]. ■

Corollary 25.5.7. *There is a natural isomorphism $\mathrm{Tor}_n^R(A, B) \cong \mathrm{Tor}_n^{R^\circ}(B, A)$.*

Proof. View the left hand side is a left Tor, and the right hand side as right Tor. They both satisfy the characteristic properties of the Tor functor and are thus isomorphic. ■

Corollary 25.5.8. *If R is abelian, then $\mathrm{Tor}_n^R(A, B) \cong \mathrm{Tor}_n^R(B, A)$.*

Therefore in the abelian case all properties of Tor are symmetric and need to be checked only for one of the arguments.

Recall that a module is called flat if the functor of the tensor product with it is exact on the corresponding side. One direction of the following theorem becomes obvious, and the converse can be proven by dimension shifting.

Theorem 25.5.9. *A right R -module F is flat iff $\mathrm{Tor}_n^R(F, _) = 0$ for all $n \geq 1$.*

- Exercise 25.1.**
1. Let $M = \prod_{i \in \mathbb{N}} \mathbb{Z}$ be the \mathbb{Z} -module of integral sequences, and e^n be the standard basis. Let $f \in \text{Hom}_{\mathbb{C}}(M, \mathbb{Z})$. Show that $f(e^n) = 0$ for almost all $n \in \mathbb{N}$, and that if $f(e^n) = 0$ for all n , then $f = 0$.
 2. Show that the \mathbb{Z} -module $\text{Hom}_{\mathbb{Z}}(M, \mathbb{Z})$ is free, and is isomorphic to $\oplus_{i \in \mathbb{N}} \mathbb{Z}$.
 3. Use the exercises above to show that the \mathbb{Z} -module M is not free, and conclude that the product of free modules need not be free.
 4. Show that the \mathbb{Z} -module \mathbb{Q} is flat.
 5. Show that the \mathbb{Z} -module \mathbb{Q} is not projective, and hence not free.
 6. Show that the \mathbb{Z} -modules \mathbb{Q}/\mathbb{Z} and \mathbb{Z}_2 are both not flat.
 7. Let R be a principal ideal domain (that is, a ring in which every ideal is generated by a single element; every element of a PID has a unique prime factorization, and any two elements have a GCD), and let M be an R -module. Show that M is injective iff M is divisible (i.e. multiplication by any $r \in R$ that is not a zero divisor is a surjective map on M). Conclude that the \mathbb{Z} -modules \mathbb{Q} and \mathbb{Q}/\mathbb{Z} are injective.
 8. Verify that the functor $\text{Hom}(_, \mathbb{Z})$ is not exact by applying it to the short exact sequence $0 \rightarrow \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \rightarrow \mathbb{Z}_2 \rightarrow 0$.
 9. Show that

$$\text{Tor}_k^{\mathbb{Z}}(\mathbb{Z}_m, \mathbb{Z}_n) = \begin{cases} \mathbb{Z}_{\gcd(n,m)}, & k = 0, 1, \\ 0, & k \geq 2. \end{cases} \quad (25.5.15)$$

Example 25.5.10. Let $R = \mathbb{Z}[t]/(t^n - 1)$ and view \mathbb{Z} as an \mathbb{R} -module where t acts as multiplication on \mathbb{Z} by 1. Show that

$$\text{Tor}_k^R(\mathbb{Z}, \mathbb{Z}) = \begin{cases} \mathbb{Z}, & k = 0, \\ \mathbb{Z}/n\mathbb{Z}, & k \text{ odd}, \\ 0, & k \text{ even}. \end{cases} \quad (25.5.16)$$

First we show that there is a free resolution

$$\dots \rightarrow R \xrightarrow{d_1} R \xrightarrow{d_0} R \xrightarrow{\pi} \mathbb{Z} \rightarrow 0, \quad (25.5.17)$$

where d_n is multiplication by $(t - 1)$ when $n > 0$ is even and by $(t^n - 1)/(t - 1)$ when n is odd. Finally, $\pi(r) = r \cdot 1$. The key to finding this resolution is the realization that R is isomorphic to the set of polynomials of degree no higher than $n - 1$. Indeed, any higher monomial t^d , $d > n$ can be uniquely rewritten as $t^{d-n} \cdot (t^n - 1)$ plus lower terms, and so in the quotient $\mathbb{Z}[t]/(t^n - 1)$ any power $d \geq n$ will vanish. Moreover this representation of each polynomial is unique, so $\ker \pi = (t - 1)R$ and this suggests that we can take $d_0(p) = (t - 1)p$. But then obviously $(t^n - 1)/(t - 1) \cdot R \subset \ker d_0$ and conversely every polynomial in $\ker d_0$ can be uniquely represented as a product of another polynomial with $(t^n - 1)/(t - 1)$. Thus

$\ker d_0 = (t^n - 1)/(t - 1) \cdot R$. This in turn suggests $d_1(p) = (t - 1)p$ so that $\text{im } d_1 = \ker d_0$. Similarly one verifies that $\ker d_1 = (t - 1)R$ and thus we can choose $d_2 = d_0$ and repeat this forever.

Since $R \otimes_R M \cong M$ for any R -module M , when we apply the functor $\underline{\otimes}_R \mathbb{Z}$ to this resolution (dropping the last term) we get:

$$\dots \rightarrow \mathbb{Z} \xrightarrow{\delta_n} \dots \rightarrow \mathbb{Z} \xrightarrow{\delta_0} \mathbb{Z}, \quad (25.5.18)$$

where $\delta_k = d_k \otimes 1_{\mathbb{Z}} : R \otimes_R \mathbb{Z} \rightarrow R \otimes_R \mathbb{Z}$. Under the isomorphism $\mathbb{Z} \rightarrow R \otimes_R \mathbb{Z}$ which is given by $k \mapsto 1 \otimes k$, $\delta_k : \mathbb{Z} \rightarrow \mathbb{Z}$ becomes the map $k \mapsto (t - 1) \cdot k = k - k = 0$ for even $k > 0$ and $k \mapsto (t^n - 1)/(t - 1) \cdot k = (t^{n-1} + t^{n-2} + \dots + 1) \cdot k = nk$ for odd k . So we can rewrite the complex as

$$\dots \rightarrow \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \xrightarrow{0} \mathbb{Z}. \quad (25.5.19)$$

So for even $k > 0$ we have $\text{Tor}_k^R(\mathbb{Z}, \mathbb{Z}) = \ker \delta_k / \text{im } \delta_{k+1} = \ker \delta_k = 0$ since multiplication by n is injective. Meanwhile for odd k , we have $\ker \delta_{k+1} = n\mathbb{Z}$ and $\ker \delta_k = \mathbb{Z}$, so $\text{Tor}_k^R(\mathbb{Z}, \mathbb{Z}) = \mathbb{Z}/n\mathbb{Z}$. For $k = 0$, we have as expected $\text{Tor}_0^R(\mathbb{Z}, \mathbb{Z}) = \mathbb{Z}/\text{im } \delta_1 = \mathbb{Z}$.

25.6 Ext functors

Unlike the Tor functor, which is entirely symmetric with respect to its two arguments, the two versions of the Ext functor will behave in very different (dual) ways. Here one can work equivalently with the category of left or right modules.

Definition 25.6.1 (Covariant Ext functor). Let A be a left (or right) R -module and consider the left exact functor $\text{Hom}_R(A, \underline{}) : R\text{-Mod} \rightarrow \mathbf{Ab}$, and define

$$\text{Ext}_R^n(A, \underline{}) = R^n(\text{Hom}_R(A, \underline{})). \quad (25.6.1)$$

Thus if $\mathbf{E} = 0 \rightarrow B \rightarrow E^0 \xrightarrow{d^0} E^1 \xrightarrow{d^1} \dots$ is an injective resolution of another left (resp. right) R -module B , then

$$\text{Ext}_R^n(A, B) = H^n(\text{Hom}(A, \mathbf{E}^B)) = \frac{\ker d_*^n}{\text{im } d_*^{n-1}}, \quad (25.6.2)$$

where $d_*^n : \text{Hom}_R(A, E^n) \rightarrow \text{Hom}_R(A, E^{n+1})$ is defined, as usual, by $d_*^n : f \mapsto d^n \circ f$.

The following characterization theorem for these functors can once again be obtained via inductive dimension shifting.

Theorem 25.6.2. Fix an R -module M . The following three properties define the covariant Ext functors $\text{Ext}_R^n(M, \underline{})$ uniquely up to natural isomorphism:

1. $\text{Ext}_R^0(M, \underline{}) \cong \text{Hom}_R(M, \underline{})$.
2. Every short exact sequence $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ gives rise to a long exact sequence

$$0 \rightarrow \text{Hom}_R(M, A) \rightarrow \text{Hom}_R(M, B) \rightarrow \text{Hom}_R(M, C) \xrightarrow{\delta^0} \text{Ext}_R^1(M, A) \rightarrow \cdots \quad (25.6.3)$$

3. If B is injective, then $\text{Ext}_R^n(M, B) = 0$, $n \geq 1$.

Definition 25.6.3 (Contravariant Ext functor). Let B be a left (or right) R -module and consider the left exact functor $\text{Hom}_R(\underline{}, B) : R\text{-Mod} \rightarrow \text{Ab}$, and define

$$\text{ext}_R^n(\underline{}, B) = R^n(\text{Hom}_R(\underline{}, B)). \quad (25.6.4)$$

Thus if $\mathbf{P} = \cdots \rightarrow P_2 \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} A \rightarrow 0$ is a projective resolution of another left (resp. right) R -module B , then

$$\text{Ext}_R^n(A, B) = H^n(\text{Hom}(\mathbf{P}_A, B)) = \frac{\ker d_n^*}{\text{im } d_{n-1}^*}, \quad (25.6.5)$$

where $d_n^* : \text{Hom}_R(P_{n-1}, B) \rightarrow \text{Hom}_R(P_n, B)$ is defined, as usual, by $d_n^* : f \mapsto f \circ d_n$.

Similarly, we have the following axiomatic characterization.

Theorem 25.6.4. Fix an R -module M . The following three properties define the contravariant Ext functors $\text{ext}_R^n(\underline{}, M)$ uniquely up to natural isomorphism:

1. $\text{ext}_R^0(\underline{}, M) \cong \text{Hom}_R(\underline{}, M)$.
2. Every short exact sequence $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ gives rise to a long exact sequence

$$0 \rightarrow \text{Hom}_R(C, M) \rightarrow \text{Hom}_R(B, M) \rightarrow \text{Hom}_R(A, M) \xrightarrow{\delta^0} \text{ext}_R^1(C, M) \rightarrow \cdots \quad (25.6.6)$$

3. If A is projective, then $\text{ext}_R^n(A, M) = 0$, $n \geq 1$.

By the same Theorem 25.5.6, we conclude the natural isomorphisms

$$\text{Ext}_R^n(A, B) \cong \text{ext}_R^n(A, B), \quad (25.6.7)$$

which means that from now on we can simplify the notation by always writing Ext_R^n regardless of whether we view it as a covariant or contravariant functor. Nevertheless, it is important to remember the duality between these two functors, e.g. when writing out their long exact sequences.

Example 25.6.5 (Continuation of Example 25.5.10). Let $R = \mathbb{Z}[t]/(t^n - 1)$ and view \mathbb{Z} as an \mathbb{R} -module where t acts as multiplication on \mathbb{Z} by 1. We will now show that

$$\mathrm{Ext}_k^R(\mathbb{Z}, \mathbb{Z}) = \begin{cases} \mathbb{Z}, & k = 0, \\ \mathbb{Z}/n\mathbb{Z}, & k > 0 \text{ even}, \\ 0, & k \text{ odd}. \end{cases} \quad (25.6.8)$$

We again take the same free resolution

$$\cdots \rightarrow R \xrightarrow{d_1} R \xrightarrow{d_0} R \xrightarrow{\pi} \mathbb{Z} \rightarrow 0, \quad (25.6.9)$$

where d_n is multiplication by $(t - 1)$ when $n > 0$ is even and by $(t^n - 1)/(t - 1)$ when n is odd. Applying the contravariant functor $\mathrm{Hom}_R(_, \mathbb{Z})$ to it and using the fact that $\mathrm{Hom}_R(R, M) \cong M$ for any R -module M , we get

$$\mathbb{Z} \xrightarrow{\delta^1} \mathbb{Z} \xrightarrow{\delta^2} \mathbb{Z} \rightarrow \cdots. \quad (25.6.10)$$

Via the natural isomorphism $\mathrm{Hom}_R(R, M) \rightarrow M$ sending f to $f(1)$, the maps $\delta^k : \mathbb{Z} \mapsto f \circ d_n$ once again become $\delta^k = 0$ for odd k and $\delta^k = n$ for even n . So the complex becomes

$$\mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \xrightarrow{0} \cdots. \quad (25.6.11)$$

When $k > 0$ is even, we have $\mathrm{Ext}_k^R(\mathbb{Z}, \mathbb{Z}) = \ker \delta^{k+1} / \mathrm{im} \delta^k = \mathbb{Z}/n\mathbb{Z}$. When k is odd, we get zero, and when $k = 0$, we have $\mathrm{Ext}_k^R(\mathbb{Z}, \mathbb{Z}) = \ker \delta^1 = \mathbb{Z}$.

26 Homological Algebra IV: Applications

26.1 Extensions

Unlike with Tor functors, it is much easier to give an explicit interpretation of the Ext functors. In fact, what we describe below is the original definition of the first Ext functor.

Definition 26.1.1 (Extension). A short exact sequence of R -modules

$$0 \rightarrow B \xrightarrow{i} E \xrightarrow{\pi} A \rightarrow 0 \quad (26.1.1)$$

is called an extension of A by B .

An extension is called *trivial* if it splits:

$$0 \rightarrow B \xrightarrow[\pi_2]{i_2} A \oplus B \xrightarrow[i_1]{\pi_1} A \rightarrow 0. \quad (26.1.2)$$

A morphism between two extensions of A by B is a morphism in the category of such short exact sequences. Since A and B are fixed, by the 5-lemma every morphism of extensions is an isomorphism. Such extensions are called *equivalent*. The set of equivalence classes of extensions of A by B is denoted $\text{Ext}(A, B)$.

We will now show that $\text{Ext}(A, B)$ in fact coincides with $\text{Ext}_R^1(A, B)$. We do this in two steps: first proving functoriality, and then establishing an isomorphism of these two objects as abelian groups.

Theorem 26.1.2. *The set $\text{Ext}(A, B)$ of equivalence classes of extensions is functorial: contravariant in A and covariant in B .*

Proof. For morphisms $f : C \rightarrow A$ and $g : B \rightarrow D$, we want to guarantee the existence of morphisms $\text{Ext}(A, B) \rightarrow \text{Ext}(C, B)$ and $\text{Ext}(A, B) \rightarrow \text{Ext}(A, D)$ that make the corresponding squares commute. In other words, we want to fill in the following diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & B & \longrightarrow & E & \longrightarrow & A & \longrightarrow 0 \\ & & \downarrow g & & & & f \uparrow & \\ 0 & \longrightarrow & D & \dashrightarrow & ? & \dashrightarrow & C & \longrightarrow 0 \end{array} \quad (26.1.3)$$

in such a way that the bottom row is an extension whose equivalence class depends only on the equivalence class of the top extension (this condition is represented by a 3D diagram).

We will split this task into two parts, building an extension of C by B to show functoriality with respect to A , and then similarly for the other side. Therefore we want to build a morphism $\text{Ext}(A, B) \rightarrow \text{Ext}(C, B)$ which will be called $\text{Ext}(f, 1_B)$. Recall from basic category theory that functoriality with respect to the first argument is often related to pullbacks (fibered products). Therefore the following diagram commutes

$$\begin{array}{ccccccc} 0 & \longrightarrow & B & \xrightarrow{i} & E & \longrightarrow & A & \longrightarrow 0 \\ & & \uparrow = & & \uparrow \pi_1 & & f \uparrow & \\ 0 & \longrightarrow & B & \xrightarrow{i \times 0} & E \times_A C & \xrightarrow{\pi_2} & C & \longrightarrow 0 \end{array} \quad (26.1.4)$$

and it is easy to verify that the bottom row is exact and hence an extension. To check functoriality with respect to the second argument, we apply a dual argument turning the left part of the diagram into the pushout square for $E \sqcup_B D$:

$$\begin{array}{ccccccc} 0 & \longrightarrow & B & \longrightarrow & E & \xrightarrow{\pi} & A & \longrightarrow 0 \\ & & \downarrow g & & \downarrow i_1 & & \downarrow = & \\ 0 & \longrightarrow & D & \xrightarrow{i_2} & E \sqcup_B D & \xrightarrow{\pi \sqcup 0} & A & \longrightarrow 0. \end{array} \quad (26.1.5)$$

Finally, one checks that equivalent extensions in the top row give rise to equivalent extensions in the bottom row by drawing the corresponding 3D diagrams and verifying their commutativity. ■

Our next goal is to find a natural group structure on $\text{Ext}(A, B)$. Obviously we can take direct sums of two different extensions of A by B in the form $0 \rightarrow B \oplus B \rightarrow E_1 \oplus E_2 \rightarrow A \oplus A \rightarrow 0$, but what we need is an extension $0 \rightarrow B \rightarrow E_1 \oplus E_2 \rightarrow A \rightarrow 0$. The natural way to obtain an extension of this form would be to apply the Ext functor which we just established to some two natural morphisms $A \rightarrow A \oplus A$ and $B \oplus B \rightarrow B$. This will let us fill in this diagram:

$$\begin{array}{ccccccc} 0 & \longrightarrow & B \oplus B & \longrightarrow & E_1 \oplus E_2 & \longrightarrow & A \oplus A & \longrightarrow 0 \\ & & \downarrow \nabla_B & & & & \Delta_A \uparrow & \\ 0 & \longrightarrow & B & \dashrightarrow & ? & \dashrightarrow & A & \longrightarrow 0. \end{array} \quad (26.1.6)$$

It is easy to guess which two maps to take:

$$\nabla_B : B \oplus B \rightarrow B, \quad (b_1, b_2) \mapsto b_1 + b_2 \quad (26.1.7)$$

$$\Delta_A : A \rightarrow A \oplus A, \quad a \mapsto (a, a). \quad (26.1.8)$$

Definition 26.1.3 (Baire sum). The sum of two extensions E_1 and E_2 of A by B is defined as

$$E_1 + E_2 = \text{Ext}(\Delta_A, \nabla_B)(E_1 \oplus E_2). \quad (26.1.9)$$

This definition of course is too abstract, so let us unpack this construction explicitly. First we build the following commuting diagram to get rid of $A \oplus A$:

$$\begin{array}{ccccccc} 0 & \longrightarrow & B \oplus B & \xrightarrow{i_1 \oplus i_2} & E_1 \oplus E_2 & \xrightarrow{\pi_1 \oplus \pi_2} & A \oplus A & \longrightarrow 0 \\ & & \downarrow = & & \uparrow & & \Delta_A \uparrow & \\ 0 & \longrightarrow & B \oplus B & \longrightarrow & (E_1 \oplus E_2) \times_{A \oplus A} A & \longrightarrow & A & \longrightarrow 0, \end{array} \quad (26.1.10)$$

where

$$\begin{aligned} (E_1 \oplus E_2) \times_{A \oplus A} A &= \{((e_1, e_2), a) \mid \pi_1(e_1) = \pi_2(e_2) = a\} = \\ &= \{(e_1, e_2) \mid \pi_1(e_1) = \pi_2(e_2)\} = E_1 \times_A E_2 \end{aligned} \quad (26.1.11)$$

since labeling the value of a is redundant after we run over all possible values of $a \in A$.

Now we continue by attaching a third row at the bottom to similarly get rid of $B \oplus B$:

$$\begin{array}{ccccccc} 0 & \longrightarrow & B \oplus B & \longrightarrow & E_1 \times_A E_2 & \longrightarrow & A \longrightarrow 0 \\ & & \downarrow \nabla_B & & \downarrow & & \downarrow = \\ 0 & \longrightarrow & B & \longrightarrow & (E_1 \times_A E_2) \sqcup_{B \oplus B} B & \longrightarrow & A \longrightarrow 0, \end{array} \quad (26.1.12)$$

where now

$$(E_1 \times_A E_2) \sqcup_{B \oplus B} B = \{(e_1, e_2) \mid \pi_1(e_1) = \pi_2(e_2)\} / \{i_1(b) - i_2(b) \mid b \in B\}. \quad (26.1.13)$$

The following becomes a very long exercise in diagram chasing.

Proposition 26.1.4 (Baire sum defines a group). *The equivalence class of $E_1 + E_2$ depends only on the equivalence classes of E_1 and E_2 and thus defines a binary operation on $\text{Ext}(A, B)$. Further, the Baire sum is commutative, associative, every equivalence class has an inverse with respect to it, and the class of trivial extensions is the neutral element.*

Thus $\text{Ext}(A, B)$ is indeed an abelian group, and we can finally establish its isomorphism with $\text{Ext}_R^1(A, B)$.

Theorem 26.1.5. *There is a natural isomorphism $\text{Ext}(_, _)$ $\cong \text{Ext}_R^1(_, _)$.*

Proof. Recall that $\text{Ext}_R^1(A, B) = H^1(\text{Hom}_R(\mathbf{P}_A, B))$. Therefore we have a diagram

$$\begin{array}{ccccccc} \cdots & \longrightarrow & P_2 & \xrightarrow{d_2} & P_1 & \xrightarrow{d_1} & P_0 \longrightarrow A \longrightarrow 0 \\ & & \downarrow 0 & & \downarrow \sigma_1 & & \downarrow \\ & & 0 & \longrightarrow & B & \longrightarrow & E \longrightarrow A \longrightarrow 0. \end{array} \quad (26.1.14)$$

By the Comparison Theorem 25.1.14, the identity morphism on A extends to a pair of dashed morphisms on P_0 and P_1 that make this diagram commute. Note that these extensions are not unique.

We need to check two properties to identify σ_1 with an element of $H^1(\text{Hom}_R(\mathbf{P}_A, B))$. First, due to the commutativity of the square on the left, $d_2^*(\sigma_1) = \sigma_1 \circ d_2 = 0$, so σ_1 is a cocycle. Next, we need to check that σ_1 is unique up to coboundaries. Let there be two different extensions by the Comparison Theorem. We know that any two such extensions are homotopic via some homotopy $h_0 : P_0 \rightarrow B$:

$$\sigma_1 - \tilde{\sigma}_1 = h_0 \circ d_1 = d_1^*(h_0), \quad (26.1.15)$$

which means that $\sigma_1 - \tilde{\sigma}_1$ is a coboundary and the class $[\sigma_1]$ is well defined.

We claim that the map $E \mapsto [\sigma_1]$ is the required isomorphism. Thus we need to check that $[\sigma_1]$ depends only on the equivalence class $[E]$. Indeed, if we've constructed an extension of 1_A to all of \mathbf{P}_A that corresponds to E_1 , then the same extension will be valid for any isomorphic extension E_2 , and conversely any two extensions are homotopic, so they will produce the same class $[\sigma_1]$.

Finally, we need to check that $[E] \mapsto [\sigma_1]$ is an homomorphism of abelian groups, i.e. $[E_1 + E_2]$ gets mapped to the sum of the images of $[E_1]$ and $[E_2]$. This is an exercise in chasing on a 3D diagram.

Now we need to construct an inverse homomorphism. This is a construction via functoriality. Since the resolution \mathbf{P} is exact, $\text{im } d_2 = \ker d_1$ and so we can quotient P_1 by this image, obtaining a short exact sequence $0 \rightarrow P_1 / \ker d_1 \rightarrow P_0 \rightarrow A \rightarrow 0$. We end up with the diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & P_1 & \xrightarrow{d_2(P_2)} & P_0 & \longrightarrow & A & \longrightarrow 0 \\ & & \downarrow & & \downarrow & & \downarrow = & \\ 0 & \longrightarrow & B & \dashrightarrow & ? & \dashrightarrow & A & \longrightarrow 0. \end{array} \quad (26.1.16)$$

The missing part of this diagram can get recovered by applying the functor Ext , obtaining a pushout object in the bottom row. We then need to check that the equivalence class of the resulting extension depends only on the cohomology class of the vertical dashed morphism, that it's a homomorphism, and that it is the inverse of the first homomorphism on both sides. ■

26.2 (*) Higher extensions

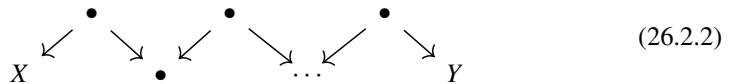
Definition 26.2.1 (Extension of length n). An extension of A by B of length n is an exact sequence of the form

$$0 \rightarrow B \rightarrow X_n \rightarrow X_{n-1} \rightarrow \cdots \rightarrow X_2 \rightarrow X_1 \rightarrow A \rightarrow 0. \quad (26.2.1)$$

A morphism between two extensions of A by B of length $n > 1$ is a morphism in the category of these sequences with 1_A and 1_B acting at the edges. Since these diagrams are too long for the 5-lemma, these morphisms are not always isomorphisms, and are called *elementary equivalences of extensions*.

Definition 26.2.2 (Yoneda equivalence of extensions). The Yoneda equivalence of extensions of A by B of length n is the coarsest (smallest) symmetric equivalence relation that extends the elementary equivalences.

This means that, naively speaking, two extensions X and Y are Yoneda equivalent iff there is a finite chain of elementary equivalences of the following shape (each dot is an extension)



The nontrivial fact about extensions is that two extensions are Yoneda equivalent if and only if there exists a single third extension that admits elementary equivalences into them, i.e. the following diagram is all one needs to look for:



Exercise 26.1. Prove the above characterization of Yoneda equivalence.

Definition 26.2.3. The set of equivalence classes of extensions of A by B of length n under the Yoneda equivalence is called $\text{Ext}^n(A, B)$.

Definition 26.2.4 (Yoneda sum). The Yoneda sum of two extensions $0 \rightarrow B \rightarrow X_n \rightarrow \dots \rightarrow X_1 \rightarrow A \rightarrow 0$ and $0 \rightarrow B \rightarrow Y_n \rightarrow \dots \rightarrow Y_1 \rightarrow A \rightarrow 0$ is

$$0 \rightarrow B \rightarrow X_n \sqcup_B Y_n \rightarrow X_{n-1} \oplus Y_{n-1} \rightarrow \dots \rightarrow X_2 \oplus Y_2 \rightarrow X_1 \times_A Y_1 \rightarrow A \rightarrow 0. \quad (26.2.4)$$

This is clearly well defined as an extension because we don't need to worry about anything more than one arrow away from A and B . More difficult is checking that the Yoneda sum respects Yoneda equivalences. We state the following without proof.

Theorem 26.2.5. *The Yoneda sum turns $\text{Ext}^n(A, B)$ into an abelian group that is naturally isomorphic to $\text{Ext}_R^n(A, B)$.*

Definition 26.2.6 (Yoneda product). The Yoneda product is a binary operation acting $\text{Ext}_R^m(B, C) \times \text{Ext}_R^n(A, B) \rightarrow \text{Ext}_R^{m+n}(A, C)$ defined for two extensions $[X] \in \text{Ext}_R^m(B, C)$ and $[Y] \in \text{Ext}_R^n(A, B)$ as their concatenation

$$0 \rightarrow C \rightarrow X_m \rightarrow \dots \rightarrow X_1 \rightarrow B \rightarrow Y_n \rightarrow \dots \rightarrow Y_1 \rightarrow A \rightarrow 0. \quad (26.2.5)$$

One can check that this product is in fact well defined on equivalence classes of extensions, and it is obviously associative.

26.3 Elementary examples

Let $R = \mathbb{Z}$, i.e. we are working with abelian groups. We will thus omit \mathbb{Z} in the notation for the functors. Let A be an abelian group. We can simultaneously compute $\text{Tor}_1(\mathbb{Z}_n, A)$ and

$\text{Hom}(\mathbb{Z}_n, A)$. First we note that in the exact sequence

$$0 \rightarrow \mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z} \rightarrow 0, \quad (26.3.1)$$

the first \mathbb{Z} is free, and hence projective and flat, which makes this a projective resolution. We can thus write the “long” exact sequence in homology:

$$\text{Tor}_1(\mathbb{Z}, A) = 0 \rightarrow \text{Tor}_1(\mathbb{Z}_n, A) \rightarrow \mathbb{Z} \otimes A \xrightarrow{\cdot n} \mathbb{Z} \otimes A \rightarrow \mathbb{Z}_n \otimes A \rightarrow 0. \quad (26.3.2)$$

We have canonically $\mathbb{Z} \otimes A = A$. The kernel of the multiplication by n is the n -torsion subgroup $A[n] = \{x \in A \mid nx = 0\}$. We have thus computed

$$\boxed{\mathbb{Z}_n \otimes A = A/nA, \quad \text{Tor}_1(\mathbb{Z}_n, A) = A[n].} \quad (26.3.3)$$

Similarly for Ext , we have the long exact sequence

$$\text{Hom}(\mathbb{Z}_n, A) \rightarrow \text{Hom}(\mathbb{Z}, A) \rightarrow \text{Hom}(\mathbb{Z}, A) \rightarrow \text{Ext}_{\mathbb{Z}}^1(\mathbb{Z}_n, A) \rightarrow \text{Ext}_{\mathbb{Z}}^1(\mathbb{Z}, A) = 0l \quad (26.3.4)$$

where we have the canonical identification $\text{Hom}(\mathbb{Z}, A) = A$, and the arrow between these two copies of A is still multiplication by n . The kernel of this map is A/nA , hence we get the dual result

$$\boxed{\text{Hom}(\mathbb{Z}_n, A) = A[n], \quad \text{Ext}^1(\mathbb{Z}_n, A) = A/nA.} \quad (26.3.5)$$

Corollary 26.3.1. $\mathbb{Z}_m \otimes \mathbb{Z}_n = \mathbb{Z}_{\gcd(m,n)}$ and $\text{Tor}_1(\mathbb{Z}_m, \mathbb{Z}_n) = \mathbb{Z}_{\gcd(m,n)}$. More generally, for two ideals $I, J \trianglelefteq R$, we have

$$R/I \otimes_R R/J = R/(I+J), \quad \text{Tor}_1^R(R/I, R/J) = I \cap J/(IJ). \quad (26.3.6)$$

The former follows from the latter because $\text{lcm}(m, n) \cdot \gcd(m, n) = mn$.

This allows us to compute Tor functors for all finitely generated groups.

Theorem 26.3.2. For abelian groups,

$$\text{Tor}_n(\bigoplus_i A_i, B) = \bigoplus_i \text{Tor}_n(A_i, B) \quad (26.3.7)$$

and similarly with respect to the second argument.

Proof. We prove by dimension shifting. Let $n = 0$, then the result is obvious since tensor products naturally distribute over direct sums. Next, representing A_i as a quotient of a projective module

$$0 \rightarrow C_i \rightarrow P_i \rightarrow A_i \rightarrow 0, \quad (26.3.8)$$

we have the exact sequence

$$0 \rightarrow \oplus_i C_i \rightarrow \oplus_i P_i \rightarrow \oplus_i A_i \rightarrow 0. \quad (26.3.9)$$

Applying Tor to this, we get the long exact sequence

$$\text{Tor}_1(\oplus_i P_i, B) = 0 \rightarrow \text{Tor}_1(\oplus_i A_i, B) \rightarrow (\oplus_i C_i) \otimes B \rightarrow (\oplus_i P_i) \otimes B \rightarrow (\oplus_i A_i) \otimes B \rightarrow 0, \quad (26.3.10)$$

where the first term vanishes since $\oplus_i P_i$ is projective. The kernel of the central arrow $\oplus_i C_i \otimes B \rightarrow \oplus_i P_i \otimes B$ is the direct sum of the kernels on each term, which is $\oplus_i \text{Tor}_1(A_i, B)$. For $n \geq 2$, we will have the later segments of the exact sequence

$$\text{Tor}_n(\oplus_i P_i, B) = 0 \rightarrow \text{Tor}_n(\oplus_i A_i, B) \rightarrow \text{Tor}_{n-1}(\oplus_i C_i, B) \rightarrow \text{Tor}_{n-1}(\oplus_i P_i, B) = 0, \quad (26.3.11)$$

therefore the central arrow is an isomorphism and since we assume by induction that $\text{Tor}_{n-1}(\oplus_i C_i, B) \cong \oplus_i \text{Tor}_{n-1}(A_i, B)$, the same property holds for Tor_n . ■

Corollary 26.3.3. *For any abelian group B ,*

$$\text{Tor}_1(\mathbb{Z}, B) = 0, \quad \text{Tor}_1(\mathbb{Z}_n, B) = B[n], \quad \text{Tor}_n(\mathbb{Z}, B) = \text{Tor}_n(\mathbb{Z}_n, B) = 0, \quad (26.3.12)$$

and hence for any finitely generated abelian group A we can apply the above theorem to find $\text{Tor}_n(A, B)$ (vanishing for $n \geq 2$).

Proof. $0 \rightarrow \mathbb{Z} \rightarrow \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z} \rightarrow 0$ is a projective resolution concentrated in degrees 0 and 1 and hence the homologies will be concentrated in those degrees as well. Tor_1 is already computed. ■

The only extra fact we used in proving that $\text{Tor}_n(A, B) = 0$ for A finitely generated and $n \geq 2$ was that $\oplus_i P_i$ is projective and hence $\text{Tor}_n(\oplus_i P_i, B) = 0$. However for this to hold it is sufficient for $\oplus_i P_i$ to be flat. Since direct limits of flat modules are flat, we can replace direct sums with arbitrary direct limits.

Theorem 26.3.4. *In the category of abelian groups, $\text{Tor}_n(\varinjlim A_i, B) = \varinjlim \text{Tor}_n(A_i, B)$.*

Any abelian group is an inductive limit of finitely generated ones, hence:

Corollary 26.3.5. *For abelian groups A and B ,*

$$\text{Tor}_n(A, B) = 0, \quad n \geq 2. \quad (26.3.13)$$

Corollary 26.3.6. *Given a short exact sequence of abelian groups $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$, there is an exact sequence*

$$0 \rightarrow \text{Tor}_1(A, D) \rightarrow \text{Tor}_1(B, D) \rightarrow \text{Tor}_1(C, D) \rightarrow A \otimes D \rightarrow B \otimes D \rightarrow C \otimes D \rightarrow 0. \quad (26.3.14)$$

This fact was unsurprisingly discovered long before the functors Tor_n^R were invented, so in the theory of abelian groups the notation is simply $\text{Tor} = \text{Tor}_1$, and can be explicitly realized via the following presentation:

$$\begin{aligned} \text{Tor}(A, B) &= \\ &= \left\langle \langle a, n, b \rangle \in A \times \mathbb{Z} \times B, \text{ s.t. } an = nb = 0 \right\rangle \begin{cases} \langle a_1 + a_2, n, b \rangle = \langle a_1, n, b \rangle + \langle a_2, n, b \rangle, \\ \langle a, n, b_1 + b_2 \rangle = \langle a, n, b_1 \rangle + \langle a, n, b_2 \rangle, \\ \langle a, mn, b \rangle = \langle am, n, b \rangle \text{ if } nb = 0, \\ \langle a, mn, b \rangle = \langle a, m, nb \rangle \text{ if } am = 0 \end{cases}. \end{aligned} \quad (26.3.15)$$

26.4 Torsion submodules

Recall that an element $a \in A$ of an R -module is called a torsion element if $n \cdot a = 0$ for some n . If R is abelian, then the set of all torsion elements forms a submodule of A , denoted $T(A)$ and called the torsion submodule. Now let R be an integral domain (a commutative ring where the product of nonzero elements is always nonzero) and let $K = Q(R)$ be its field of fractions. In this case $A/T(A)$ is a torsion-free submodule. Any torsion-free R -module can be embedded in a K -vector space:

$$0 \rightarrow A \rightarrow A \otimes_R K \rightarrow B \rightarrow 0, \quad B = T(A), \quad (26.4.1)$$

where B is a torsion module.

Theorem 26.4.1. *Let R be an integral domain. Then for any R -module A , $T(A) \cong \text{Tor}_1(K/R, A)$.*

For the proof we will use the following three lemmas.

Lemma 26.4.2. *If $T(A) = A$, then $\text{Tor}_1(K/R, A) \cong A$.*

Proof. Consider the exact sequence

$$0 \rightarrow R \rightarrow K \rightarrow K/R \rightarrow 0 \quad (26.4.2)$$

and look at the end of the corresponding Tor sequence:

$$\mathrm{Tor}_1(K, A) \rightarrow \mathrm{Tor}_1(K/R, A) \rightarrow A \otimes_R R = A \rightarrow A \otimes_R K. \quad (26.4.3)$$

Now we notice that K is a flat R -module (as is any localization of a ring), so $\mathrm{Tor}_1(K, A) = 0$. Moreover, K is divisible, which means that $K \otimes_R A = 0$ (we can always write $k \otimes a = \frac{k}{n} \otimes na = 0$ if $na = 0$). The exactness of the sequence then implies the asserted isomorphism. ■

Lemma 26.4.3. $\mathrm{Tor}_n(K/R, A) = 0$ for all $n \geq 2$.

Proof. This time we have the exact segment at $n \geq 2$:

$$\mathrm{Tor}_n(K, A) \rightarrow \mathrm{Tor}_n(K/R, A) \xrightarrow{\delta} \mathrm{Tor}_{n-1}(R, A). \quad (26.4.4)$$

Here we notice that R is projective and hence $\mathrm{Tor}_{n-1}(R, A) = 0$, whereas K is flat and thus $\mathrm{Tor}_n(K, A) = 0$. Exactness implies the assertion. ■

Lemma 26.4.4. If $T(A) = 0$, then $\mathrm{Tor}_1(K/R, A) = 0$.

Proof. Finally we have the remaining exact segment

$$\mathrm{Tor}_2(K/R, B) \rightarrow \mathrm{Tor}_1(K/R, A) \rightarrow \mathrm{Tor}_1(K/R, A \otimes_R K). \quad (26.4.5)$$

Here we notice that $A \otimes_R K$ is flat (a vector space) since K is flat, hence $\mathrm{Tor}_1(K/R, A \otimes_R K) = 0$. Also $\mathrm{Tor}_2(K/R, B) = 0$ by Lemma 26.4.3 since B is a torsion module. ■

Proof of Theorem 26.4.1. We have the long exact sequence of Tor functors:

$$\mathrm{Tor}_2(K/R, A/T(A)) \rightarrow \mathrm{Tor}_1(K/R, T(A)) \rightarrow \mathrm{Tor}_1(K/R, A) \rightarrow \mathrm{Tor}_1(K/R, A/T(A)). \quad (26.4.6)$$

The two groups at the edges are trivial by Lemma 26.4.3 and Lemma 26.4.4, respectively. Moreover $\mathrm{Tor}_1(K/R, T(A)) = T(A)$ by Lemma 26.4.2. The exactness then implies the asserted isomorphism. ■

Higher Tor functors don't have such a simple interpretation, but we can at least show that all of them are still torsion modules.

Theorem 26.4.5. Let R be an integral domain, let $n \geq 1$, and let A, B be R -modules. Then $\mathrm{Tor}_n(A, B)$ is a torsion module.

Proof. First we prove the special case when B is a torsion module, i.e. $B = T(B)$, and in this case we can even allow $n = 0$: for $n = 0$ we get $A \otimes B$, which is a torsion module because $n(a \otimes b) = a \otimes nb$. For $n \geq 1$, we can perform dimension shifting. Let P be a projective module giving an exact sequence of the form (e.g. it can be the free module generated by the generators of A)

$$0 \rightarrow C \rightarrow P \rightarrow A \rightarrow 0. \quad (26.4.7)$$

Now we have the exact segment

$$\mathrm{Tor}_n(P, B) \rightarrow \mathrm{Tor}_n(A, B) \rightarrow \mathrm{Tor}_{n-1}(C, B). \quad (26.4.8)$$

The left end vanishes because P is projective. Moreover, $\mathrm{Tor}_{n-1}(C, B)$ is a torsion module by the assumption of the induction step. Therefore $\mathrm{Tor}_n(A, B)$ is embedded into a torsion module and hence must be a torsion module itself.

Now we consider another special case where B is torsion-free, i.e. $T(B) = 0$. We can build another exact sequence of the form

$$0 \rightarrow B \rightarrow B \otimes_R K \rightarrow C \rightarrow 0, \quad (26.4.9)$$

and thus its exact sequence reads, in part,

$$\mathrm{Tor}_{n+1}(A, C) \rightarrow \mathrm{Tor}_n(A, B) \rightarrow \mathrm{Tor}_1(A, B \otimes_R K). \quad (26.4.10)$$

But the right end vanishes because $B \otimes_R K$ is flat (a vector space) and $\mathrm{Tor}_{n+1}(A, C)$ is a torsion module since C is. Therefore $\mathrm{Tor}_n(A, B)$ is the image of a torsion module and hence must be a torsion module itself.

Now we can deal with the general case. We consider the exact sequence

$$0 \rightarrow T(B) \rightarrow B \rightarrow B/T(B) \rightarrow 0 \quad (26.4.11)$$

and get the corresponding exact sequence

$$\mathrm{Tor}_n(A, T(B)) \rightarrow \mathrm{Tor}_n(A, B) \rightarrow \mathrm{Tor}_n(A, B/T(B)). \quad (26.4.12)$$

We have just shown that the left and right ends of this sequence are torsion modules, hence $\mathrm{Tor}_n(A, B)$ is as well (by virtue of being an extension of a torsion module by another torsion module). ■

26.5 (*) Application in class field theory

Obviously $\mathrm{Ext}_{\mathbb{Z}}^1(\mathbb{Z}, \mathbb{Q}) = 0$ (either because \mathbb{Z} is projective or because \mathbb{Q} is injective) but after reversing the arguments the result becomes highly nontrivial, and in fact the cardinality of continuum.

Theorem 26.5.1. *Ext functors respect the following direct sums (coproducts) and direct products for finite and infinite index sets:*

$$\mathrm{Ext}_R^n(\oplus_i A_i, B) = \prod_i \mathrm{Ext}_R^n(A_i, B), \quad \mathrm{Ext}_R^n(A, \prod_i B_i) = \prod_i \mathrm{Ext}_R^n(A, B_i). \quad (26.5.1)$$

Proof. Similar proof by dimension shifting, using the fact that direct sums of projectives are projective and products of injectives are injective. ■

However we quickly start to run into issues because, unlike direct sums (which are an example of a direct limit), other direct limits of injective modules are not always injective. Generally, limits of injectives/projectives are not injective/projective, so Tor and Ext functors commute only with finite limits. If they did commute with infinite limits, since $\mathbb{Q} = \varinjlim \frac{1}{n}\mathbb{Z}$, we would get a trivial $\mathrm{Ext}^1(\mathbb{Q}, \mathbb{Z})$ since $\mathrm{Ext}^1(\mathbb{Z}, \mathbb{Z}) = 0$. We present this computation purely “for fun”, but it is actually foundational in several areas of modern mathematics, such as the Langlands program and the theory of algebraic groups.

Definition 26.5.2 (Group of adeles). Recall that the p -adic integers can be defined as $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}$. Denote their field of fractions by $\mathbb{Q}_p = Q(\mathbb{Z}_p)$. The additive group of finite adèles of \mathbb{Q} is defined as the *restricted* product of \mathbb{Q}_p over all prime numbers p , i.e. almost all components must land in \mathbb{Z}_p :

$$\mathbb{A}_f(\mathbb{Q}) = \prod'_{p \in \mathbb{Z}_p} \mathbb{Q}_p = \{(x_2, x_3, x_5, \dots) \mid x_p \in \mathbb{Q}_p \text{ and almost all } x_p \in \mathbb{Z}_p\}. \quad (26.5.2)$$

Theorem 26.5.3. $\mathrm{Ext}_{\mathbb{Z}}^1(\mathbb{Q}, \mathbb{Z}) \cong \mathbb{A}_f(\mathbb{Q})/\mathbb{Q}$.

Proof. The foundation of this fact is the short exact sequence

$$0 \rightarrow \mathbb{Z} \rightarrow \mathbb{Q} \rightarrow \mathbb{Q}/\mathbb{Z} \rightarrow 0, \quad (26.5.3)$$

which is at once a flat resolution of \mathbb{Q}/\mathbb{Z} and an injective resolution of \mathbb{Z} . We can take the corresponding long exact sequence for the functor $\mathrm{Ext}(\mathbb{Q}, -)$:

$$0 \rightarrow \mathrm{Hom}(\mathbb{Q}, \mathbb{Z}) \rightarrow \mathrm{Hom}(\mathbb{Q}, \mathbb{Q}) \rightarrow \mathrm{Hom}(\mathbb{Q}, \mathbb{Q}/\mathbb{Z}) \rightarrow \mathrm{Ext}^1(\mathbb{Q}, \mathbb{Z}) \rightarrow \mathrm{Ext}^1(\mathbb{Q}, \mathbb{Q}). \quad (26.5.4)$$

We know that $\mathrm{Ext}^1(\mathbb{Q}, \mathbb{Q}) = 0$ since \mathbb{Q} is an injective \mathbb{Z} -module. Now, $\mathrm{Hom}(\mathbb{Q}, \mathbb{Z}) = 0$: if $q \in \mathbb{Q}$, then we can write $q = nq'$ for any n , so if $f \in \mathrm{Hom}(\mathbb{Q}, \mathbb{Z})$ then $f(q) = nf(q')$, i.e. $f(q)$ is divisible by n for all n , which is possible only if $f = 0$. We also know

$\text{Hom}(\mathbb{Q}, \mathbb{Q}) \cong \mathbb{Q}$ since any \mathbb{Z} -linear map from \mathbb{Q} to \mathbb{Q} is just multiplication by an element of \mathbb{Q} . So the exact sequence implies the isomorphism

$$\text{Ext}^1(\mathbb{Q}, \mathbb{Z}) \cong \text{Hom}(\mathbb{Q}, \mathbb{Q}/\mathbb{Z})/\mathbb{Q}. \quad (26.5.5)$$

Thus it suffices to describe the group $\text{Hom}(\mathbb{Q}, \mathbb{Q}/\mathbb{Z})$.

Let's start by describing the structure of \mathbb{Q}/\mathbb{Z} . For any prime p , there is the subgroup $\mathbb{Z}/p^\infty\mathbb{Z} = \mathbb{Z}[p^{-1}]/\mathbb{Z}$, called the (additive) Prüfer group³⁸, consisting of elements of the form a/p^k with $p \nmid a$ and $0 \leq a \leq p^k$. Putting all these subgroups together, we get a map from $\bigoplus_p \mathbb{Z}[p^{-1}]/\mathbb{Z} = \bigoplus_p (\mathbb{Z}[p^{-1}]/\mathbb{Z})$ into \mathbb{Q}/\mathbb{Z} . This map is injective: if $\frac{a}{m} + \frac{b}{n} = \frac{an+bm}{nm} = 0$ in \mathbb{Q}/\mathbb{Z} with n, m coprime, then $\frac{an+bm}{nm} \in \mathbb{Z}$, i.e. $nm \mid an + bm$, so $n \mid bm$ and $m \mid an$. But since n, m are coprime, this means $m \mid a$ and $n \mid b$. Thus $\frac{a}{m}$ and $\frac{b}{n}$ are in \mathbb{Z} , so they are 0 in \mathbb{Q}/\mathbb{Z} . Now, we can write an element of $\bigoplus_p \mathbb{Z}[p^{-1}]/\mathbb{Z}$ as:

$$\frac{a_1}{p_1^{k_1}} + \cdots + \frac{a_n}{p_n^{k_n}} = \frac{N}{p_1^{k_1} \cdots p_{n-1}^{k_{n-1}}} + \frac{a_n}{p_n^{k_n}}. \quad (26.5.6)$$

Thus, the above argument shows that $\frac{a_n}{p_n^{k_n}} \in \mathbb{Z}$, so we may perform induction on n to show that the whole sum is in \mathbb{Z} and therefore vanishes in the quotient $\bigoplus_p \mathbb{Z}[p^{-1}]/\mathbb{Z}$.

Now let us show that this map is also surjective. To do this, let $\frac{a}{q} \in \mathbb{Q}$ with $q = q_1 q_2$ coprime. Then we may write $1 = aq_1 + bq_2$ for some $a, b \in \mathbb{Z}$ (e.g. by the Chinese Remainder Theorem, or by the fact that \mathbb{Z} is a principal integral domain, so the ideal (q_1, q_2) is $(\gcd(q_1, q_2)) = (1)$). Then we can take the “partial fraction” decomposition:

$$\frac{1}{q_1 q_2} = \frac{aq_1 + bq_2}{q_1 q_2} = \frac{a}{q_2} + \frac{b}{q_1}. \quad (26.5.7)$$

By breaking q into its prime factorization and repeatedly using this identity, we may write q as an element of the image of $\bigoplus_p \mathbb{Z}[p^{-1}]/\mathbb{Z}$. This proves the isomorphism

$$\mathbb{Q}/\mathbb{Z} \cong \bigoplus_{p \in \mathbb{P}} \mathbb{Z}[p^{-1}]/\mathbb{Z}. \quad (26.5.8)$$

As an aside, we can notice that

$$\mathbb{Z}[p^{-1}]/\mathbb{Z} \cong \mathbb{Z}/p^\infty\mathbb{Z} = \varinjlim \mathbb{Z}/p^n\mathbb{Z}, \quad (26.5.9)$$

³⁸It is isomorphic to (multiplicative) cyclic group of roots of unity $C_{p^\infty} = \bigcup_{n>0} C_{p^n} \cong \{\epsilon \in \mathbb{C}^k \mid \exists n \in \mathbb{N} : \epsilon^{p^n} = 1\}$. Here, $\mathbb{Z}[p^{-1}]$ is the additive group of all rational numbers whose denominators are powers of p .

where the direct limit is with respect to the embeddings $\mathbb{Z}/p^n\mathbb{Z} \xrightarrow{\cdot^n} \mathbb{Z}/p^{n+1}\mathbb{Z}$. This is contrasted with the p -adic integers $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}$, which is a projective limit taken with respect to the epimorphisms given by $\mathbb{Z}/p^n\mathbb{Z} \xrightarrow{\text{mod } p^{n-1}} \mathbb{Z}/p^{n-1}\mathbb{Z}$. The fundamental difference between these two limits is that $\text{mod } n$ is a homomorphism of rings with unity, whereas $\cdot n$ is not. This means that the projective limit defining \mathbb{Z}_p will preserve the multiplicative structure of a ring, whereas $\mathbb{Z}/p^\infty\mathbb{Z}$ supports only an additive structure.

A direct sum $\bigoplus_i G_i$ of an infinite sequence of groups G_i consists of sequences (g_i) which are almost everywhere identity. This is strictly a subset of the direct product $\prod_i G_i$, which consists of all possible sequences (g_i) . By the universal property of products, a map to the product is the same as a tuple of maps to each factor, i.e.

$$\text{Hom}\left(\mathbb{Q}, \prod_p \mathbb{Z}/p^\infty\mathbb{Z}\right) = \prod_p \text{Hom}(\mathbb{Q}, \mathbb{Z}/p^\infty\mathbb{Z}). \quad (26.5.10)$$

We will deal with the difference between products and sums later. For now the next step is to compute $\text{Hom}(\mathbb{Q}, \mathbb{Z}/p^\infty\mathbb{Z})$. Such homomorphisms of course restrict to homomorphisms from $\mathbb{Z}[p^{-1}]$ to $\mathbb{Z}[p^{-1}]/\mathbb{Z}$, and in fact any such homomorphism extends *uniquely* to \mathbb{Q} . To see this, we will use the following claims.

Claim 1. The group $G = \mathbb{Z}/p^\infty\mathbb{Z}$ is *uniquely divisible* by numbers coprime to p : for any $\alpha \in G$ and $n \in \mathbb{N}$ with $p \nmid n$, there is a unique $\alpha' \in G$ such that $n \cdot \alpha' = 0$.

Proof of Claim 1. We can write $\alpha = \frac{a}{p^k} + \mathbb{Z}$ with $p \nmid a$, $0 \leq a < p^k$. Since $p \nmid n$, n and p^k are coprime, so there are $b, c \in \mathbb{Z}$ with $bp^k + cn = 1$, so $abp^k + acn = a$. Thus we can write α as

$$\alpha = \frac{a}{p^k} + \mathbb{Z} = \frac{abp^k}{p^k} + \frac{acn}{p^k} + \mathbb{Z} = n \cdot \frac{ac}{p^k} + \mathbb{Z}. \quad (26.5.11)$$

So we may take $\alpha' = \frac{ac}{p^k} + \mathbb{Z}$. We need to show that α' is unique, so let β' be an element of G with $n\beta' = \alpha$. Then $n(\beta' - \alpha') = 0$, so it suffices to show that multiplication by n is injective on G . Now, let $\gamma = \frac{m}{p^l} + \mathbb{Z} \in G$. If $n\gamma = 0$ then $\frac{nm}{p^l} \in \mathbb{Z}$, so $p^l \mid nm$. Since $p \nmid n$, this implies $p^l \mid m$, so $\gamma = 0$. ■

Now, let $f \in \text{Hom}(\mathbb{Z}[p^{-1}], G)$. We want to show that it extends uniquely to $\tilde{f} \in \text{Hom}(\mathbb{Q}, G)$. Write any element of \mathbb{Q} uniquely as $\frac{a}{p^k m}$ with $p \nmid m$, $\gcd(p^k m, a) = 1$, and $m > 0$. Let $\alpha = f\left(\frac{a}{p^k}\right)$, which is defined since $a \in \mathbb{Z}[p^{-1}]$. We can define $\tilde{f}\left(\frac{a}{p^k m}\right)$ as the unique element α' such that $m \cdot \alpha' = \alpha$. This gives a well-defined function from \mathbb{Q} to $G = \mathbb{Z}[p^{-1}]/\mathbb{Z}$,

and it is easy to see that it is additive and extends f . Moreover, it is unique since we must have $m \cdot \tilde{f}\left(\frac{a}{p^k m}\right) = f\left(\frac{a}{p^k}\right)$.

Thus we need to determine $\text{Hom}(\mathbb{Z}[p^{-1}], \mathbb{Z}[p^{-1}]/\mathbb{Z})$. Keeping the notation $G = \mathbb{Z}[p^{-1}]/\mathbb{Z}$, let f_0 be such a homomorphism and consider $\alpha = f_0(1) \in G$. We have $\alpha = \frac{a}{p^n}$ for some $n \geq 0$ with $p \nmid a$, so $p^n \cdot \alpha = a = 0$ and $p^m a \neq 0$ for $m < n$. Define $f = p^n f_0$, so $f(1) = 0$. Then define a sequence $m_n = f\left(\frac{1}{p^n}\right)$ for $n \geq 1$. We have $p \cdot m_n = m_{n-1}$, and $p^n m_n = f(1) = 0$ for all n . On the other hand, given any sequence (m_n) with $m_n \in G$ with these two properties, we can define a homomorphism $f : \mathbb{Z}[p^{-1}] \rightarrow G$ such that $m_n = f\left(\frac{1}{p^n}\right)$ and $f(1) = 0$. To do this, define $f\left(\frac{a}{p^n}\right) = am_n$ with $a \in \mathbb{Z}$. If we rewrite $\frac{a}{p^n}$ as $\frac{ap}{p^{n+1}}$, then since $apm_{n+1} = am_n$, these definitions agree. This allows us to check that f is additive:

$$f\left(\frac{a}{p^n} + \frac{b}{p^k}\right) = f\left(\frac{ap^k + bp^n}{p^{n+k}}\right) = ap^k m_{k+n} + bp^n m_{k+n} = am_n + bm_k = f\left(\frac{a}{p^n}\right) + f\left(\frac{b}{p^k}\right). \quad (26.5.12)$$

Therefore, to describe the set $\text{Hom}(\mathbb{Z}[p^{-1}], G)$, it suffices describe the set of sequences $(m_n) \in G$ with the two properties $p \cdot m_n = m_{n-1}$ and $p^n m_n = 0$. The second property says that $m_n = \frac{a}{p^n}$ for some a (perhaps not coprime with p). Since a is only defined $\pmod{p^n}$, we can think of m_n as an element of $\mathbb{Z}/p^n\mathbb{Z}$ instead. Then $pm_n = \frac{pa}{p^{n+1}} = \frac{a}{p^{n-1}}$, so the first property can be rephrased as saying that $m_n = m_{n-1} \pmod{p^{n-1}}$ in $\mathbb{Z}/p^n\mathbb{Z}$. Thus the subgroup of $\text{Hom}(\mathbb{Z}[p^{-1}], G)$ consisting of maps f with $f(1) = 0$ is isomorphic to the group of sequences (m_n) with $m_n \in \mathbb{Z}/p^n\mathbb{Z}$ such that $\pi_{n,n-1}(m_n) = m_{n-1}$, where $\pi_{n,n-1}(x) = x \pmod{p^{n-1}}$. This group is exactly the group of p -adic integers \mathbb{Z}_p .³⁹

Now, for any $f \in \text{Hom}(\mathbb{Z}[p^{-1}], G)$ an any n , there is a unique $f_0 \in \text{Hom}(\mathbb{Z}[p^{-1}], G)$ with $p^n f_0 = f$: we can take $f_0(x) = f\left(\frac{x}{p^n}\right)$, and this is unique since multiplication by p^n is injective on $\mathbb{Z}[p^{-1}]$. Thus, $\text{Hom}(\mathbb{Z}[p^{-1}], G)$ has a unique structure of a $\mathbb{Z}[p^{-1}]$ -module. Since for any $f \in \text{Hom}(\mathbb{Z}[p^{-1}], G)$, there is some n such that $p^n f(1) = 0$, we can write f as f_1/p^n with $f_1(1) = 0$. This shows that $\text{Hom}(\mathbb{Z}[p^{-1}], G) \cong \mathbb{Z}_p[p^{-1}] = \mathbb{Q}_p$, the p -adic numbers as an abelian group.

³⁹For an abelian group M , note that $d \mid n$ implies $nM < dM$, so there is a quotient map $\pi_n^d : M/mM \rightarrow M/dM$ (it descends to identity on M , so in symbols it's just $\bar{m} \mapsto \bar{m}$). Now define the submodule $\text{consist}(M)$ of $\prod_{n \in \mathbb{N}} M/nM$ by $\text{consist}(M) = \{(m_n \in M/nM)_{n \in \mathbb{N}} \mid d \mid n \Rightarrow \pi_n^d(m_n) = m_d\}$. Then $\mathbb{Z}_p = \text{consist}(\mathbb{Z}[p^{-1}])$, where since $p^k m \cdot \mathbb{Z}[p^{-1}] = p^k \mathbb{Z}[p^{-1}]$ for $p \nmid m$.

Another way of seeing this is to notice

$$\text{Hom}(G, G) = \text{End}(G) = \text{End}\left(\varprojlim \mathbb{Z}/p^n\mathbb{Z}\right) = \varprojlim \text{End}(\mathbb{Z}/p^n\mathbb{Z}) = \varprojlim \mathbb{Z}/p^n\mathbb{Z} = \mathbb{Z}_p, \quad (26.5.13)$$

where we used the fact that $\text{End}(C_k) = C_k$ for any cyclic group C_k . Next, any $f \in \text{Hom}(\mathbb{Z}[p^{-1}], G)$ must satisfy $p^n f(1) = 0$ for some n , so f becomes an element of $\text{Hom}(G, G)$ after being multiplied by some power of p , and is thus identified with an element of \mathbb{Z}_p . Therefore

$$\text{Hom}(\mathbb{Z}[p^{-1}], G) = \varinjlim p^{-n}\mathbb{Z}_p = \mathbb{Z}_p[p^{-1}] = \mathbb{Q}_p. \quad (26.5.14)$$

In summary, we have

$$\text{Hom}(\mathbb{Q}, \mathbb{Q}/\mathbb{Z}) \cong \text{Hom}\left(\mathbb{Q}, \bigoplus_p \mathbb{Z}[p^{-1}]/\mathbb{Z}\right) < \text{Hom}\left(\mathbb{Q}, \prod_p \mathbb{Z}[p^{-1}]/\mathbb{Z}\right) \cong \prod_p \mathbb{Q}_p. \quad (26.5.15)$$

The difference between the direct sum and the direct product is that in the direct sum, almost all components must be identity, so in the Hom-sets this turns into the condition that the homomorphism is represented by a sequence of maps (f_p) such that $f_p(x) = 0$ for almost all prime p . We claim that this is exactly the subgroup of $\prod_p \mathbb{Q}_p$ consisting of sequences (a_p) such that $a_p \in \mathbb{Z}_p$ for almost all p . To see this, let $x = \frac{m}{n}$. For all $p \nmid nm$, $f_p(x) = a \cdot f_p(1)$ for some $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$, by definition of the isomorphism $\text{Hom}(\mathbb{Z}[p^{-1}], \mathbb{Z}[p^{-1}]/\mathbb{Z}) \cong \text{Hom}(\mathbb{Q}, \mathbb{Z}[p^{-1}]/\mathbb{Z})$ and the proof of Claim 1. Thus $f_p(x) = 0$ iff $f_p(1) = 0$. Therefore a sequence (f_p) satisfies the condition that for all x , $f_p(x) = 0$ for almost all p if and only if $f_p(1) = 0$ for almost all p , which is equivalent to requiring that the corresponding elements $a_p \in \mathbb{Q}_p$ belong to $\mathbb{Z}_p \subset \mathbb{Q}_p$ for almost all p . We obtain the group

$$\mathbb{A}_f(\mathbb{Q}) = \prod'_{\mathbb{Z}_p} \mathbb{Q}_p, \quad (26.5.16)$$

where $\prod'_{\mathbb{Z}_p}$ stands for “restricted product” and indicates the subset of the direct product where almost all entries are in \mathbb{Z}_p . This group has a natural ring structure by component-wise multiplication and is called the *finite adèle ring of \mathbb{Q}* .⁴⁰

Finally, we see that $\text{Ext}^1(\mathbb{Q}, \mathbb{Z}) \cong \mathbb{A}_f(\mathbb{Q})/\mathbb{Q}$, where the map $\mathbb{Q} \rightarrow \mathbb{F}$ is given by sending $q \in \mathbb{Q}$ to the maps $(f_p) : \mathbb{Q} \rightarrow \mathbb{Z}[p^{-1}]/\mathbb{Z}$ defined as multiplication by q for each p . This corresponds to an element $(\iota_p(q)) \in \mathbb{A}_f(\mathbb{Q})$ with $\iota_p : \mathbb{Q} \rightarrow \mathbb{Q}_p$ defined by sending $\frac{a}{p^k m}$ with

⁴⁰The full adèle ring $\mathbb{A}(\mathbb{Q})$ is $\mathbb{A}_f(\mathbb{Q}) \times \mathbb{R}$: sometimes it is useful to think of \mathbb{R} as “the prime at infinity”. This ring has a locally compact topology coming from the locally compact topologies on \mathbb{R} and \mathbb{Q}_p , and many important results in number theory can be reformulated in terms of this topological ring.

$p \nmid a, m$ to $\frac{1}{p^k} (m^{-1} \cdot a \pmod p)^n$ (where m^{-1} is an inverse to $a \pmod{p^n}$, which exists for each n but depends on n). Since the denominator of q is only divisible by finitely many primes, we see that $\iota_p(q) \in \mathbb{Z}_p$ for almost all p , so this in fact lands in $\mathbb{A}_f(\mathbb{Q})$. ■

26.6 Homology with coefficients

Suppose that we are given integer homology (e.g. singular homology) $H_*(X)$. Clearly instead of having the coefficients of simplices in chains be integers, they could be elements of any ring R . This way we can define $H_*(X; R)$. It satisfies all Eilenberg-Steenrod axioms with the only change that the zeroth homology of a singleton space is $H_0(\{\}; R) = R$. Consider, for example, $H_*(\mathbb{R}P^n; \mathbb{Z}_2)$. The cellular chain complex has $C_k(\mathbb{R}P^n; \mathbb{Z}_2) = \mathbb{Z}_2$ for $0 \leq k \neq n$, and the boundary operator alternates between multiplication by 2 and by 0 (see formula 24.11.34). But in \mathbb{Z}_2 , $2 = 0$, therefore $\partial = 0$ and the homology mimics the chain complex itself:

$$H_k(\mathbb{R}P^n; \mathbb{Z}_2) = \begin{cases} \mathbb{Z}_2, & 0 \leq k \leq n, \\ 0, & k > n. \end{cases} \quad (26.6.1)$$

On the other hand, suppose R is a ring in which 2 is invertible, such as F_p for p odd. This time $\partial_{2k+1} = 0$ and ∂_{2k} is an isomorphism if n is even, and vice versa if n is odd. Therefore

$$H_k(\mathbb{R}P^{2n}; R) = \begin{cases} R, & k = 0, \\ 0, & k \neq 0, \end{cases} \quad H_k(\mathbb{R}P^{2n+1}; R) = \begin{cases} R, & k = 0, n, \\ 0, & k \neq 0, n. \end{cases} \quad (26.6.2)$$

We get a much simpler result: away from 2, projective spaces look like points, and odd projective spaces look like spheres.

To pose a more general question: knowing $H_*(X) = H_*(X; \mathbb{Z})$, can we compute, for example, $H_*(X; \mathbb{Z}_2)$? One could be tempted to find a map of topological spaces that induces an isomorphism of these homologies. Consider the map $\mathbb{R}P^2 \rightarrow \mathbb{S}^2$ that pinches $\mathbb{R}P^1 \subset \mathbb{R}P^2$ to a point. Since $H_2(\mathbb{R}P^2; \mathbb{Z}) = 0$, this map induces a zero map in H_2 . But in \mathbb{Z}_2 -coefficients, in dimension 2, this map induces an isomorphism. This shows that there is no *natural* (functorial) determination of $H_*(X; \mathbb{Z}_2)$ in terms of $H_*(X; \mathbb{Z})$: the effect of a map in integer homology does not determine its effect in homology mod 2.

Even more generally, given $H_*(X; R)$ for some ring R , can we compute $H_*(X; M)$ for some R -module M ? For this we will need to first define the latter object.

We will now restrict to the case when R is a principal ideal domain (PID), although there are generalizations for other rings as well. Let (\mathbf{C}, d) be a chain complex of R -modules. Given another R -module G , we can construct the chain complex $\mathbf{C} \otimes G = (C_n \otimes_R G, d_n \otimes 1)$. First we compute the homologies of this complex in the case of free modules.

Definition 26.6.1 (Homology with coefficients). Given a complex of R -modules (\mathbf{C}, d) and an R -module G , the homology groups of \mathbf{C} with coefficients in G , denoted $H_n(\mathbf{C}, G)$, are the homology groups of the complex $C_\bullet \otimes G$.

The universal coefficients theorem states that the homology with coefficients in any G are completely determined by integer homology (which is why it's called universal).

Proposition 26.6.2 (Universal coefficients in homology [26, Prop. 11.9.1]). *Let (\mathbf{C}, d) be a chain complex of free modules. Then there exists an exact sequence*

$$0 \rightarrow H_n(\mathbb{C}) \otimes G \xrightarrow{\alpha} H_n(\mathbf{C}, G) \xrightarrow{\partial} \text{Tor}(H_{n-1}(\mathbf{C}), G) \rightarrow 0. \quad (26.6.3)$$

The sequence is natural in \mathbf{C} and G and splits, thereby identifying $H_n(\mathbf{C}, G)$ with a direct sum. The homomorphism α sends $[z] \otimes g$ for a cycle z to the homology class $[z \otimes g]$.

Proof. The sequence $0 \rightarrow Z_n \rightarrow C_n \xrightarrow{d_n} B_{n-1} \rightarrow 0$ is exact; B_{n-1} is a submodule of C_{n-1} and hence free. Therefore the sequence splits, making $Z_n \rightarrow C_n$ a split monomorphism, and the induced sequence

$$0 \rightarrow Z_n \otimes G \rightarrow C_n \otimes G \rightarrow B_{n-1} \otimes G \rightarrow 0 \quad (26.6.4)$$

is again a split exact sequence, making $Z_n \otimes G \rightarrow C_n \otimes G$ a split monomorphism. The rest of the work is to show that this induces a splitting of α in homology.

We consider the totality of these sequences as an exact sequence of chain complexes, where the Z - and the B -complex have trivial boundary operators. Associated to this short exact sequence of chain complexes is a long exact sequence in homology

$$B_n \otimes G \xrightarrow{i \otimes 1} Z_n \otimes G \rightarrow H_n(\mathbf{C} \otimes G) \rightarrow B_{n-1} \otimes G \xrightarrow{i \otimes 1} Z_{n-1} \otimes G. \quad (26.6.5)$$

One verifies that the connecting homomorphism is exactly $i \otimes 1$, where $i : B_n \hookrightarrow Z_n$. The sequence $B_n \otimes G \rightarrow Z_n \otimes G \rightarrow H_n \otimes G \rightarrow 0$ is exact, hence the cokernel of $i \otimes 1$ is $H_n(\mathbf{C}) \otimes G$, and the resulting map $H_n(\mathbf{C}) \otimes G \rightarrow H_n(\mathbf{C} \otimes G)$ is α . The kernel of $B_{n-1} \otimes G \rightarrow Z_{n-1} \otimes G$ is $\text{Tor}(H_{n-1}(\mathbf{C}), G)$, because $0 \rightarrow B_{n-1} \rightarrow Z_{n-1} \rightarrow H_{n-1}(\mathbf{C}) \rightarrow 0$ is a free resolution. Let $r : C_n \rightarrow Z_n$ be a splitting of $Z_n \hookrightarrow C_n$. Then

$$Z_n(\mathbf{C} \otimes G) \hookrightarrow C_n \otimes G \xrightarrow{r \otimes 1} Z_n \otimes G \rightarrow H_n(\mathbf{C}) \otimes G \quad (26.6.6)$$

maps $B_n(\mathbf{C} \otimes G)$ to zero and induces $\rho : H_n(\mathbf{C} \otimes G) \rightarrow H_n(\mathbf{C}) \otimes G$ with $\rho \circ \alpha = 1$, i.e. a splitting of the universal coefficient sequence. ■

A dual result exists for cohomology. This definition shows that this is the point where homology and cohomology really diverge. Even though cohomology is simply homology of complexes in negative degrees, adding coefficients severs this simple connection. Although homology and cohomology are still formally dual, cohomology theories have many common tools (like ring structures) whose duals in homology are very poorly understood or are simply unconventional.

Definition 26.6.3 (Cohomology with coefficients). Given a complex of R -modules (\mathbf{C}, d) and an R -module G , the cohomology groups of \mathbf{C} with coefficients in G , denoted $H^n(\mathbf{C}, G)$, are the cohomology groups of the cocomplex $\text{Hom}(C_\bullet, G)$.

Proposition 26.6.4 (Universal coefficients in cohomology [26, Prop. 11.9.2]). *Let (\mathbf{C}, d) be a chain complex of free modules. There exists an exact sequence*

$$0 \rightarrow \text{Ext}(H_{n-1}(\mathbf{C}), G) \rightarrow H^n(\mathbf{C}, G) \xrightarrow{\alpha} \text{Hom}(H_n(\mathbf{C}), G) \rightarrow 0. \quad (26.6.7)$$

The map α sends the cohomology class of the cocycle $\varphi : C_n \rightarrow G$ to the homomorphism $H_n(\mathbf{C}) \rightarrow G$, $[c] \mapsto \varphi(c)$. The sequence is natural with respect to chain maps acting on \mathbf{C} and homomorphisms acting on G . It splits, therefore identifying $H^n(\mathbf{C}, G)$ with a direct sum, and the splitting is natural in G but not in \mathbf{C} .

Proof. Again we start with the split exact sequence $0 \rightarrow Z_n \rightarrow C_n \rightarrow B_{n-1} \rightarrow 0$ and the induced exact sequence

$$0 \rightarrow \text{Hom}(B_{n-1}, G) \rightarrow \text{Hom}(C_n, G) \rightarrow \text{Hom}(Z_n, G) \rightarrow 0. \quad (26.6.8)$$

We consider the totality of these sequences as an exact sequence of cochain complexes with the Z - and B -complex having trivial coboundary operators. Associated to this short exact sequence of cochain complexes is a long exact sequence in cohomology

$$\cdots \rightarrow \text{Hom}(B_{n-1}, G) \rightarrow H^n(\mathbf{C}, G) \rightarrow \text{Hom}(Z_n, G) \xrightarrow{d^n} \text{Hom}(B_n, G) \rightarrow \cdots \quad (26.6.9)$$

which induces a short exact sequence

$$0 \rightarrow \text{coker } d^{n-1} \rightarrow H^n(\mathbf{C}, G) \xrightarrow{\alpha} \ker d^n \rightarrow 0. \quad (26.6.10)$$

The rest of the proof is similar to the homological case, except we first need to verify that the connecting homomorphism d^n is in fact induced by the inclusion $i : B_n \hookrightarrow Z_n$ so that it splits, which we show in the next lemma.

Lemma 26.6.5 ([26, Lem. 11.9.3]). *The formal coboundary operator d^n (without the additional sign introduced earlier!) is the homomorphism induced by $i : B_n \hookrightarrow Z_n$.*

Proof. Let $\varphi : Z_n \rightarrow G$ be given. Then $d^n(\varphi)$ is obtained as follows: extend φ to $\tilde{\varphi} : C_n \rightarrow G$. Apply δ and find a preimage of $\delta(\tilde{\varphi}) = \tilde{\varphi} \circ d_{n+1}$ in $\text{Hom}(B_n, G)$. One verifies that $\varphi \circ i$ is such a preimage. ■

From the exact sequence $0 \rightarrow B_n \rightarrow Z_n \rightarrow H_n(\mathbf{C}) \rightarrow 0$ we obtain the exact sequence

$$0 \rightarrow \text{Hom}(H_n(\mathbf{C}), G) \rightarrow \text{Hom}(Z_n, G) \xrightarrow{i^*} \text{Hom}(B_n, G). \quad (26.6.11)$$

We use it to get the isomorphism $\ker i^* \cong \text{Hom}(H_n(\mathbf{C}), G)$. One verifies that α is as claimed. From the free resolution and the definition of Ext we thus obtain the claimed exact sequence. The naturality of this sequence is a consequence of the construction. It remains to verify the splitting. Pick a splitting $r : C_n \rightarrow Z_n$ of the inclusion $Z_n \hookrightarrow C_n$ and consider the diagram

$$\begin{array}{ccccccc} 0 & \rightarrow & Z^n(\text{Hom}(\mathbf{C}, G)) & \hookrightarrow & \text{Hom}(C_n, G) & \xrightarrow{\delta} & \text{Hom}(C_{n+1}, G) \\ & & \uparrow & & r^* \uparrow & & \downarrow \\ 0 & \rightarrow & \text{Hom}(H_n(\mathbf{C}), G) & \rightarrow & \text{Hom}(Z_n, G) & \xrightarrow{i^*} & \text{Hom}(B_n, G). \end{array} \quad (26.6.12)$$

If $\varphi \in \ker i^*$, then $r^*(\varphi) = \varphi \circ r \in \ker \delta$. The splitting (dashed arrow) is then induced by the morphism $\ker i^* \rightarrow Z^n(\text{Hom}(\mathbf{C}, G))$ given by $\varphi \mapsto \varphi \circ t$. ■

26.7 Künneth formula

27 Homological Algebra V: Spectral Sequences

Use Meinrenken as guide, also Bott-Tu, etc. They will be used to prove the De Rham theorem.

28 Cohomology of Topological Spaces

28.1 Eilenberg-Zilber maps

28.2 Singular cohomology

28.3 de Rham theorem

28.4 de Rham currents

28.5 Borsuk-Ulam theorem

Follow Bredon.

28.6 Fiber sequences

28.7 Simplicial sets, classifying spaces

Yin's notes. Also May's book and Miller's notes.

28.8 Eilenberg-MacLane spaces

Use May's book and Miller's notes, Fushida-Hardy notes.

28.9 Postnikov, Whitehead towers

28.10 Representability of cohomology

29 Group (Co)Homology

29.1 Definition

Definition 29.1.1 (Group ring). Let G be a group and R a ring. The group ring $R[G]$ is defined as the set of maps $f : G \rightarrow R$ with finite support. Pointwise addition and convolution of these maps turn $R[G]$ into a ring. The elements of $R[G]$ can be represented by formal linear combinations of the form $\sum_{g \in G} r_g \cdot g$, $r_g \in R$, with finitely many nonzero terms.

To study group (co)homology we will need G to act on an abelian group A so that A becomes a $\mathbb{Z}[G]$ -module. In the most classical situations, $A = \mathbb{Z}$ and the action of G on A is trivial. For simplicity we will assume that the action is a left action, but as usual everything works for right actions too. Some of the first group (co)homologies (H^1) were introduced by Hilbert and I. Schur at the dawn of the 20th century, extended by O. Schreier and R. Brauer in the 1920s (H^2). The first general definition of $H_n(G, \mathbb{Z})$ was given by W. Hurewicz in the 1930s. In 1940 O. Teichmüller introduced $H^3(G, A)$, and over the course of the 1940s two parallel teams, Eilenberg and MacLane in the US, along with D. Faddeev in the USSR, discovered the general theory of group (co)homologies.

Definition 29.1.2 (G -module). Let A be an abelian group and G an arbitrary group. The structure of a G -module on A is a *linear* action of G on A (which can be seen as a representation of G over \mathbb{Z}). Linearity of this action is equivalent to saying that A is a $\mathbb{Z}[G]$ -module. Note that \mathbb{Z} is canonically a $\mathbb{Z}[G]$ -module with the trivial action of G .

Definition 29.1.3 (Group cohomology). Let A be an abelian group and G an arbitrary group acting on A from the left, $G \curvearrowright A$. The n -th cohomology group of G with coefficients in A is

$$H^n(G, A) = \text{Ext}_{\mathbb{Z}[G]}^n(\mathbb{Z}, A). \quad (29.1.1)$$

Note that this functor is covariant with respect to A but contravariant with respect to G (leading to some suggestions that the names of group homologies and cohomologies should have been reversed).

To interpret this object let us return to the axiomatic characterization of Ext functors.

1. $H^0(G, A) = \text{Ext}_{\mathbb{Z}[G]}^0(\mathbb{Z}, A) = \text{Hom}_{\mathbb{Z}G}(\mathbb{Z}, A)$, i.e. the set of G -equivariant morphisms $f : \mathbb{Z} \rightarrow A$. But the action on \mathbb{Z} is trivial, so the condition reduces to $f(1) = f(g \cdot 1) = g \cdot f(1)$. Therefore each such f can be identified with an element of the *module of G -invariants*

$$A^G = \{a \in A : g \cdot a = a \text{ for all } g \in G\}. \quad (29.1.2)$$

We need to confirm that $(G, A) \mapsto A^G$ is in fact a functor contravariant in G and covariant in A (acting $\text{Gr} \times G\text{-Mod} \rightarrow \text{Gr}$, and in fact into G -modules).

Let $f : A \rightarrow B$ be a homomorphism of G -modules, i.e. an equivariant linear map. Then we claim that $f(A^G) \leq B^G$. Let $a \in A^G$, $g \in G$, and write

$$g \cdot f(a) = f(g \cdot a) = f(a). \quad (29.1.3)$$

This means that to each $f : A \rightarrow B$ corresponds a $f_* : A^G \rightarrow B^G$. This confirms that A^G is covariant with respect to A .

Now let $\phi : H \rightarrow G$ be a group homomorphism. Any G -module A can be made into a H -module as follows. Let $h \in H$, $a \in A$, and define $h \cdot a = \phi(h) \cdot a$. We have $A^G \leq A^H$ because $A^H = A^{\phi(H)}$ and $\phi(H)$ is a subset of G , so it allows no fewer invariants. This confirms contravariance with respect to G .

Thus A^G is a true functor. It's easy to check that this functor is left exact: given a short exact sequence of G -modules $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$, the sequence

$$0 \rightarrow A^G \rightarrow B^G \rightarrow C^G \quad (29.1.4)$$

is also exact. We have

$$\boxed{H^0(G, A) = A^G}. \quad (29.1.5)$$

2. Given a short exact sequence of G -modules $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$, there are connecting homomorphisms that make the following long sequence exact:

$$0 \rightarrow A^G \rightarrow B^G \rightarrow C^G \rightarrow H^1(G, A) \rightarrow H^1(G, B) \rightarrow H^1(G, C) \xrightarrow{\delta} H^2(G, A) \rightarrow \dots \quad (29.1.6)$$

3. $H^n(G, P) = 0$, $n \geq 1$, if P is a projective $\mathbb{Z}[G]$ -module .

Definition 29.1.4 (Group homology). Let A be an abelian group and G an arbitrary group acting on A from the left, $G \curvearrowright A$. The n -th homology group of G with coefficients in A is

$$H_n(G, A) = \text{Tor}_n^{\mathbb{Z}[G]}(\mathbb{Z}, A). \quad (29.1.7)$$

Note that this functor is covariant with respect to A and covariant with respect to G .

Now we interpret $H_0(G, A) = \mathbb{Z} \otimes_{\mathbb{Z}[G]} A$. Since \mathbb{Z} is embedded in $\mathbb{Z}[G]$, all integer numbers can be absorbed into A and thus this set is generated by elements of the form $1 \otimes_{\mathbb{Z}[G]} a$, $a \in A$. Now, keeping in mind that the tensor product is over $\mathbb{Z}[G]$, so we can swap the action of G between the two factors, we note that

$$g \cdot (1 \otimes a) = 1 \otimes g \cdot a = g \cdot 1 \otimes a = 1 \otimes a. \quad (29.1.8)$$

Defining the augmentation map $\eta : \mathbb{Z}[G] \rightarrow \mathbb{Z}$ by $\eta(g) = 1$ for all $g \in G$ (i.e. it simply counts the number of nonzero terms in the element of the group ring), we have the exact sequence

$$0 \rightarrow I_G \rightarrow \mathbb{Z}[G] \xrightarrow{\eta} \mathbb{Z} \rightarrow 0, \quad (29.1.9)$$

where $I_G = \ker \eta$ is the *augmentation ideal*. This generates the long exact sequence

$$\cdots \rightarrow I_G \otimes_{\mathbb{Z}[G]} A \rightarrow A \rightarrow \mathbb{Z} \otimes_{\mathbb{Z}[G]} A \rightarrow 0, \quad (29.1.10)$$

where we used $\mathbb{Z}[G] \otimes_{\mathbb{Z}[G]} A = A$. Therefore $\mathbb{Z} \otimes_{\mathbb{Z}[G]} A$ is identified with the *module of coinvariants*

$$A_G = A / \langle ga - a \mid g \in G, a \in A \rangle = A / (I_G \cdot A). \quad (29.1.11)$$

We have

$$H_0(G, A) = A_G. \quad (29.1.12)$$

The other two axioms defining higher H_n 's are analogous and we don't spell them out.

To compute (co)homologies we will need to construct projective (in fact free) of \mathbb{Z} as a $\mathbb{Z}[G]$ -module. The beginning of the resolution is naturally given by the augmentation map η .

Theorem 29.1.5. I_G is a free \mathbb{Z} -module generated by $g - 1$ with $g \in G^+ = G \setminus \{1\}$.

Proof. η is a ring homomorphism and $\eta(g - 1) = 1 - 1 = 0$, thus any \mathbb{Z} -linear combination of elements of the form $g - 1$ belongs to $I_G = \ker \eta$.

Conversely, if $z \in I_G$, then $z = \sum_g n_g g$ with $\sum_g n_g = 0$. Thus

$$0 = \sum_g n_g g = \sum_g n_g g - \left(\sum_g n_g g \right) \cdot 1 = \sum_g n_g (g - 1), \quad (29.1.13)$$

where we used the definition of the group ring operations. Obviously all elements of the form $g - 1, g \in G^+$, are linearly independent over \mathbb{Z} . ■

There is one special case when I_G is not just a \mathbb{Z} -module, but a $\mathbb{Z}[G]$ -module (which requires us to be able to find a much smaller basis).

Theorem 29.1.6. *If $G = F = F(X)$ is a free group freely generated by a set X , then I_F is a free $\mathbb{Z}[F]$ -module generated by elements of the form $x - 1, x \in X$.*

Proof. For any $g, h \in G, gh - 1 = gh - g + g - 1 = g(h - 1) + (g - 1)$ and $g^{-1} - 1 = -g^{-1}(g - 1)$. This shows that, if G is generated by X (not even freely), then I_G is generated by $\{x - 1\}_{x \in X}$ over $\mathbb{Z}[G]$.

The converse will require more work. Given a map $X \rightarrow A$, we need to find the unique F -equivariant map $I_F \rightarrow A$ through which the given map factors (where the embedding $X \hookrightarrow I_F$ is $x \mapsto x - 1$). Uniqueness follows from the fact that $\{x - 1\}_x$ is a generating set. These maps are called *crossed homomorphisms*, or 1-cocycles. ■

This leads to the free resolution of free groups being very short.

Corollary 29.1.7. *For a free group F , the sequence*

$$0 \rightarrow I_F \hookrightarrow \mathbb{Z}[F] \xrightarrow{\eta} \mathbb{Z} \rightarrow 0 \quad (29.1.14)$$

is a free resolution. In particular, $H^n(F, A) = H_n(F, A) = 0$ for $n \geq 2$ (in degree $n = 1$ they may still be nontrivial, as we will see below).

29.2 Cyclic groups

Let $G = C_m$ be a cyclic group of order $m < \infty$ generated by an element x . Then there exists a “periodic resolution”. Let $D = x - 1$ and $N = 1 + x + \dots + x^{m-1} \in \mathbb{Z}[G]$. We have $DN = ND = 0$.

Theorem 29.2.1. *The complex*

$$\dots \rightarrow \mathbb{Z}[G] \xrightarrow{D} \mathbb{Z}[G] \xrightarrow{N} \mathbb{Z}[G] \xrightarrow{D} \mathbb{Z}[G] \xrightarrow{N} \mathbb{Z}[G] \xrightarrow{D} \mathbb{Z}[G] \xrightarrow{\eta} \mathbb{Z} \rightarrow 0, \quad (29.2.1)$$

where N, D denote multiplications by the elements N and D (which are G -equivariant maps since G is abelian), is a free resolution of $\mathbb{Z}[G]$ -modules.

Proof. Exactness in the last $\mathbb{Z}[G]$ is already known because it is generated by the image of D as we have just seen (because $x^i - 1$ is generated and is a multiple of $(x - 1)$). Further we need to show that $\ker D \subset \text{Im } N$ and $\ker N \subset \text{Im } D$.

Let $z = \sum_{i=0}^{m-1} a_i x^i \in \ker D$. Then $0 = \sum_{i=0}^{m-1} (a_{i+1} - a_i) x^i$, which means $a_0 = a_1 = \dots = a_{m-1}$ and $z = a_0 N$, hence $z \in \text{Im } N$.

Now suppose $z \in \ker N$, which means $0 = \eta(Nz) = \eta(N)\eta(z) = m\eta(z)$, thus $\eta(z) = 0$ and we can find a preimage:

$$z = D(-a_0 - (a_0 + a_1)x - (a_0 + a_1 + a_2)x^2 - \dots - (a_0 + \dots + a_{m-1})x^{m-1}), \quad (29.2.2)$$

where the last term actually vanishes. This proves $z \in \text{Im } D$. ■

Corollary 29.2.2. Denote

$${}_N A = \{a \in A \mid Na = 0\}. \quad (29.2.3)$$

If $G = C_m$, then $H^0(G, A) = 0$, and for $n \geq 1$, $H^{2n-1}(G, A) = {}_N A / DA$ and $H^{2n}(G, A) = A^G / NA$.

Corollary 29.2.3. If $G = C_m$, then $H_0(G, A) = A_G = A / (I_G \cdot A)$, $H_{2n-1}(G, A) = A^G / NA$ and $H_{2n}(G, A) = {}_N A / DA$ for $n \geq 1$.

In other words, for $n \geq 1$,

$$H_{2n-1}(G, A) = H^{2n}(G, A), \quad H_{2n}(G, A) = H^{2n-1}(G, A). \quad (29.2.4)$$

Remark 29.2.4. The periodicity of (co)homologies are quite rare, however the fact that homologies and cohomologies can be expressed in terms of each other is a more general fact that is called “Tate cohomology”, where both homologies and cohomologies fall into a single exact sequence (after some adjustment of H^0 and H_0 for finite groups and a more complicated story for infinite groups).

29.3 Two resolutions

Now we will construct two resolutions for an arbitrary group G .

Definition 29.3.1 (Homogeneous/Standard resolution). The following complex is called the homogeneous (or *standard*) resolution of a group G . For $n \geq 0$, let F_n be the \mathbb{Z} -module generated by the basis of n -tuples $(g_0, g_1, \dots, g_n) \in G^{n+1}$. We give F_n the structure of a G -module by

$$g(g_0, \dots, g_n) = (gg_0, \dots, gg_n). \quad (29.3.1)$$

We have $F_0 \cong \mathbb{Z}[G]$ and the augmentation map $\eta : F_0 \rightarrow \mathbb{Z}$, $(g_0) \mapsto 1$. The boundary

operator $\partial_n : F_n \rightarrow F_{n-1}$ is defined as in simplicial complexes:

$$\partial_n(g_0, \dots, g_n) = \sum_{i=0}^n (-1)^i (g_0, \dots, \hat{g}_i, \dots, g_n), \quad (29.3.2)$$

where the hat denotes an omitted argument. Just like in simplices, we get $\partial_{n-1} \circ \partial_n = 0$, which confirms that the above is a complex. Exactness is verified below.

To show that the homogeneous resolution is exact (or acyclic) it suffices to show that it is contractible. We will construct the corresponding homotopy from the identity map to the zero map. We define

$$s_{n-1} : F_{n-1} \rightarrow F_n, \quad (g_0, \dots, g_{n-1}) \mapsto (1, g_0, \dots, g_{n-1}). \quad (29.3.3)$$

Now it is trivial to check that

$$(\partial_{n+1} s_n + s_{n-1} \partial_n)(g_0, \dots, g_n) = (g_0, \dots, g_n). \quad (29.3.4)$$

The second resolution will have a much smaller basis because it's freely generated over $\mathbb{Z}[G]$ instead of \mathbb{Z} , but the price is that proving its exactness is much harder.

Definition 29.3.2 (Inhomogeneous/Bar resolution). The following complex is called the inhomogenous (or bar) resolution of a group G . For $n \geq 0$, let B_n be the $\mathbb{Z}[G]$ -modules freely generated (over $\mathbb{Z}[G]!$) by tuples $[g_1 \mid g_2 \mid \dots \mid g_n] \in G^n$. In particular B_0 is generated by the empty tuple $[] = 1$. We also have the augmentation map $\eta : B_0 \rightarrow \mathbb{Z}$ defined by $\eta(1) = 1$. The boundary operators $d_n : B_n \rightarrow B_{n-1}$ are defined by

$$\begin{aligned} d_n[g_1 \mid \dots \mid g_n] &= g_1[g_2 \mid \dots \mid g_n] + \sum_{i=2}^{i-1} (-1)^i [g_1 \mid \dots \mid g_{i-1} \mid g_i g_{i+1} \mid \dots \mid g_n] + \\ &\quad + (-1)^n [g_1 \mid \dots \mid g_{n-1}]. \end{aligned} \quad (29.3.5)$$

Exercise 29.1. Check that $d_{n-1} \circ d_n = 0$.

The result of this exercise will also follow from the following theorem, which is actually easier than doing it directly.

Theorem 29.3.3. *The homogeneous and the inhomogeneous resolutions are isomorphic. In particular, the latter is indeed a resolution.*

Proof. As a sanity check, let us apply the first few ∂_n 's to the bar resolution:

$$\partial_0([]) = 1, \quad \partial_1([g]) = g[] - [] = g - 1, \quad \partial_2([h \mid g]) = h[g] - [hg] + [h], \quad (29.3.6)$$

and finally

$$\partial_3([f \mid g \mid h]) = f[g \mid h] - [fg \mid h] + [f \mid gh] - [f \mid g], \quad (29.3.7)$$

and these are exactly the right expressions to generate F_0, F_1, F_2 , and F_3 , respectively.

Now we construct the isomorphism by providing two quasi-inverse chain maps:

$$F_n \rightarrow B_n, \quad (g_0, \dots, g_n) \mapsto g_0[g_0^{-1}g_1 \mid \dots \mid g_{n-1}^{-1}g_n], \quad (29.3.8)$$

and

$$B_n \rightarrow F_n, \quad [g_1 \mid \dots \mid g_n] \mapsto (1, g_1, g_1g_2, g_1g_2g_3, \dots, g_1g_2 \dots g_n). \quad (29.3.9)$$

It is an exercise to check that these are morphisms of chain complexes (the difficult part) and that they are mutually inverse (the easy part). ■

Definition 29.3.4 (Normalized resolutions). Denoting by F either the homogeneous or the inhomogeneous resolution of G , let $\bar{F}_* = F_*/D_*$, where D_n in the homogeneous case is generated by elements of the form (g_0, \dots, g_n) where $g_i = g_{i+1}$ for some $i \in \{0, 1, \dots, n-1\}$; and in the inhomogeneous case it's generated by $[g_1 \mid \dots \mid g_n]$ with $g_i = 1$ for some $i \in \{1, \dots, n\}$.

Theorem 29.3.5. \bar{F}_* is a free resolution of \mathbb{Z} over $\mathbb{Z}[G]$.

Proof. It suffices to prove the theorem for the homogeneous resolution, and the other will follow because it's isomorphic to the first and the two versions of D_* are mapped exactly to each other by the isomorphism of resolutions. Note that D_* is a subcomplex of F_* (exercise) that is mapped into itself by the contracting homotopy (inserting an extra 1 into the tuples doesn't break the defining properties of D_* in either case). Therefore the contracting homotopy on F_* , which we defined in the homogeneous case, will descend to \bar{F}_* , proving its contractibility. ■

29.4 H_1 and H^1

Theorem 29.4.1. $H_1(G, \mathbb{Z}) \cong G^{\text{ab}} = G/[G, G]$. Consequently, if the action $G \curvearrowright A$ is trivial, then $H_1(G, A) \cong G^{\text{ab}} \otimes_{\mathbb{Z}} A$.

Proof. Consider the sequence $0 \rightarrow I_G \rightarrow \mathbb{Z}[G] \xrightarrow{\eta} \mathbb{Z} \rightarrow 0$ and its derived sequence in homology:

$$0 = H_1(G, \mathbb{Z}[G]) \rightarrow H_1(G, \mathbb{Z}) \xrightarrow{\delta} H_0(G, I_G) \rightarrow H_0(G, \mathbb{Z}[G]) \rightarrow H_0(G, \mathbb{Z}) \rightarrow 0, \quad (29.4.1)$$

where the leftmost term vanishes because $\mathbb{Z}[G]$ is a projective $\mathbb{Z}[G]$ -module. Now, coinvariants are the largest quotient module on which G acts trivially. But the action on \mathbb{Z} is already trivial, thus $H_0(G, \mathbb{Z}) = \mathbb{Z}$. The largest quotient module of $\mathbb{Z}[G]$ on which G acts trivially is also \mathbb{Z} : $H_0(G, \mathbb{Z}[G]) = \mathbb{Z}[G]/(I_G \cdot \mathbb{Z}[G]) = \mathbb{Z}[G]/I_G = \mathbb{Z}$ (where we used the fact that I_G is an ideal so $I_G\mathbb{Z}[G] = \mathbb{Z}[G]$). This means that η descends to a surjective (by exactness) map $\eta_* : \mathbb{Z} \rightarrow \mathbb{Z}$, but all such maps are isomorphisms. Thus we have

$$0 \rightarrow H_1(G, \mathbb{Z}) \xrightarrow{\delta} H_0(G, I_G) \rightarrow \mathbb{Z} \xrightarrow{\cong} \mathbb{Z} \rightarrow 0. \quad (29.4.2)$$

The exactness of this sequence implies that δ is an isomorphism:

$$H_1(G, \mathbb{Z}) \cong H_0(G, I_G) = (I_G)_G = I_G/I_G^2. \quad (29.4.3)$$

It remains to prove that $I_G/I_G^2 \cong G/[G, G]$. This isomorphism is given by the two mutually inverse maps:

$$\alpha : G \rightarrow I_G/I_G^2, \quad g \mapsto (g - 1) + I_G^2. \quad (29.4.4)$$

We check the homomorphism property:

$$\alpha(hg) = (hg - 1) + I_G^2 = (h - 1)(g - 1) + (h - 1) + (g - 1) + I_G^2 = (h - 1) + (g - 1) + I_G^2 = \alpha(h) + \alpha(g). \quad (29.4.5)$$

Since I_G/I_G^2 is abelian, this also proves that $[G, G] \leq \ker \alpha$, so we can view this as a homomorphism $\alpha : G/[G, G] \rightarrow I_G/I_G^2$. Its inverse can be given by

$$\beta : I_G \rightarrow G/[G, G], \quad (g - 1) \mapsto g[G, G]. \quad (29.4.6)$$

We notice that $\beta(I_G^2) = 0$ thus this descends to

$$\beta : I_G/I_G^2 \rightarrow G/[G, G], \quad (g - 1) + I_G^2 \mapsto g[G, G]. \quad (29.4.7)$$

The same chain of equalities (29.4.5) confirms that this definition of β is consistent. It also shows that β is a homomorphism. It should be obvious that α and β are mutual inverses, proving the theorem. ■

Now we move on to $H^1(G, A)$ in the general case. Given a left action $G \curvearrowright A$, we have the trivial extension of G by A , called the semidirect product:

$$1 \rightarrow A \rightarrow A \rtimes G \rightarrow G \rightarrow 1. \quad (29.4.8)$$

The group $A \rtimes G$ can be realized as the Cartesian product of sets $A \times G$ with the multiplication law $(a, h)(b, g) = (a + (h \cdot b), hg)$. The intuition for this definition is clear in multiplicative

notation, where the action is $h \cdot b = hbh^{-1}$ and the pairs should be read as products, so we have $(ah)(bg) = (ahbh^{-1})(hg) = ahbg$ as expected; the conjugation arises from “commuting” h and b to write the result in the same form.

Now imagine we pick a subgroup of $A \rtimes G$ which is projected onto G , that is, pick a splitting map $G \rightarrow A \rtimes G$ for the above exact sequence. Of course there is one canonical splitting given by $g \mapsto (1, g)$. Is it true that any slitting is conjugate to the canonical one? Let the splitting be $g \mapsto (a_g, g)$. For this to be a homomorphism we need

$$(a_{hg}, hg) = (a_h, h)(a_g, g) = (a_h + h \cdot a_g, hg), \quad (29.4.9)$$

which means

$$a_{hg} = a_h + h \cdot a_g. \quad (29.4.10)$$

Definition 29.4.2 (Crossed homomorphism). Given a linear action $G \curvearrowright A$ of a group G on a G -module A , a crossed homomorphism is a map $\varphi : G \rightarrow A$ such that

$$\varphi(hg) = h \cdot \varphi(g) + \varphi(h). \quad (29.4.11)$$

The set of all such maps is denoted $\text{Der}(G, A)$ and is an abelian group.

Definition 29.4.3 (Principal crossed homomorphism). A special class of crossed homomorphisms, called principal, is parametrized by an element $a \in A$ and acts by $g \mapsto g \cdot a - a$. It is easy to check that these are crossed homomorphisms for any $a \in A$.

Therefore each splitting of a semidirect product determines a crossed homomorphism. The natural question now is: when do two different crossed homomorphisms generate two conjugate splittings? In other words, how many crossed homomorphism up to conjugation of the splittings are there?

Theorem 29.4.4. $H^1(G, A) \cong \text{Der}(G, A)/\{\text{principal crossed homs}\}$.

Proof. A 1-cocycle is a 1-chain f such that $d^1 f = 0$. We have, using the bar resolution, that this condition selects exactly the set of crossed homomorphisms:

$$0 = (d^1 f)(h, g) = h \cdot f(g) - f(g) + f(h). \quad (29.4.12)$$

A 1-coboundary is an image of d^0 , which has exactly the form $(g - 1) \cdot a$, which is a principal crossed homomorphism. The assertion follows by definition of the cohomology group. ■

Corollary 29.4.5. *The set of sections $G \rightarrow A \rtimes G$ up to conjugation in $A \rtimes G$ is naturally isomorphic to $H^1(G, A)$.*

Proof. Conjugation by an element of G doesn't change a section (which is a crossed homomorphism), whereas a conjugation by an element of A corresponds exactly to adding a principal crossed homomorphism. ■

Example 29.4.6. Consider the defining representation $\mathrm{SL}_n(K) \curvearrowright K^n$ by $(g, u) \mapsto gu$ (for some field K). Then we can ask whether $H^1(\mathrm{SL}_n(K), K^n)$ is trivial, which is equivalent to asking whether the following representation is completely reducible:

$$\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix}, \quad A \in \mathrm{SL}_n(K), v \in K^n. \quad (29.4.13)$$

In other words, is it true that every subgroup of the set of such matrices that is isomorphic to $\mathrm{SL}_n(K)$ conjugate to the canonical subgroup with $v = 0$? Surprisingly, for several finite fields the answer turns out to be no, e.g. $H^1(\mathrm{SL}_4(\mathbb{F}_2), \mathbb{F}_2^4) \neq 0$. All such cases were classified by I. Schur and Higman.

Definition 29.4.7. $\mathrm{Der}(G, A) \cong \mathrm{Hom}_G(I_G, A)$.

Proof. The homomorphism $\mathrm{Der}(G, A) \rightarrow \mathrm{Hom}_G(I_G, A)$ is

$$\phi \mapsto ((g - 1) \mapsto \phi(g)), \quad (29.4.14)$$

and its inverse is

$$f \mapsto (g \mapsto f(g - 1)). \quad (29.4.15)$$

It is an exercise to check that these are mutually inverse. We have already used this computation for proving the short free resolution of \mathbb{Z} in the case of a free group $G = F_X$, and hence all homologies of free groups in degrees at least 2 vanish. ■

29.5 H_2 and H^2

Definition 29.5.1 (Schur multiplier). The Schur multiplier of a group G is $M(G) = H_2(G, \mathbb{Z})$.

It is non-canonically isomorphic to the following object.

Theorem 29.5.2 (Hopf [27, Thm. 9.63]). *Let the group G have a given presentation, so that $G \cong F/R$, where F is a free group and R is a normal (free) subgroup generated by a set of relations. Then $H_2(G, \mathbb{Z}) \cong (R \cap [F, F])/[R, F]$.*

Proof. Exercise. On a more fundamental level this arises from the fact that $1 \rightarrow R \rightarrow F \rightarrow G \rightarrow 1$ is an exact sequence, and we could develop yet another set of derived functors with respect to this group argument, as opposed to the module argument (A). This leads to the concepts of resision, corestriction (transfer), inflation, and coinflation. ■

Historically Schur was interested in projective representations of groups, i.e. homomorphisms $\varphi : G \rightarrow \mathrm{PGL}_n(K)$. This is equivalent to replacing the homomorphism law for regular linear representations with

$$\varphi_{gh} = \lambda_{h,g} \varphi_g \varphi_h, \quad \lambda h, g \in K^*. \quad (29.5.1)$$

Associativity then forces λ to be a cocycle. Schur was studying when a projective representation can be lifted to a regular one.

A group G is called *perfect* if $G = [G, G]$, i.e. $H^1(G, \mathbb{Z}) = 0$. Topologically this is a requirement that G be “connected”. The Schur multiplier then explains the algebraic analog of being simply connected.

Definition 29.5.3 (Central extension). A group epimorphism $\pi : E \rightarrow G$ is called a *central extension* if $\ker \pi \leq Z(E)$. Morphisms of central extensions are homomorphisms $E_1 \rightarrow E_2$ that create commuting triangles. A *trivial central extension* is one isomorphic to $\mathrm{pr}_1 : G \times C \rightarrow G$, where C is an abelian group. G is called *centrally closed* if it has no nontrivial central extensions. A central extension is called *universal* if it is the terminal object in the category of central extensions of G .

Theorem 29.5.4 (Schur). *If G is perfect, then it has a universal central extension $E \xrightarrow{\pi} G$ which is a centrally closed group, and $\ker \pi = M(G)$.*

This is the discrete analog of the existence of universal covering groups and $M(G)$ is the analog of $\pi_1(G)$.

Theorem 29.5.5 (Schur). *The groups $\mathrm{SL}_n(\mathbb{F}_q)$, $n \geq 5$, are centrally closed (i.e. have trivial Schur multipliers).*

Moving on to cohomology, Schreier theory deals with extensions by abelian groups.

Definition 29.5.6 (Extension with abelian kernel). Let G be a group. An extension of G with an *abelian kernel* A is an exact sequence

$$1 \rightarrow A \rightarrow X \rightarrow G \rightarrow 1. \quad (29.5.2)$$

In an extension by an abelian kernel, any two preimages of a point $g \in G$ differ from each other by an element of G . This means that every such extension defines a consistent action of G on A by $g \cdot a = x_g a x_g^{-1}$ where $x_g \in X$ is an arbitrary preimage of g . Schreier theory classifies all extensions that generate the same action.

Theorem 29.5.7 (Schreier). *Extensions of G with an abelian kernel A with a fixed action $G \curvearrowright A$, up to isomorphism, can be identified with elements of $H^2(G, A)$. This bijection can be turned into an isomorphism of groups.*

Proof. Choosing an arbitrary set of preimages x_g , we have $x_hx_h = \sigma(h, g)x_{hg}$, where $\sigma : G \times G \rightarrow X$ is called a *factor set*. Now expand the associativity law $(x_fx_h)x_g = x_f(x_hx_g)$ and conclude that

$$\sigma(f, h)\sigma(fh, g) = f\sigma(h, g)f^{-1}\sigma(f, hg). \quad (29.5.3)$$

In the additive notation this reads

$$f \cdot \sigma(h, g) - \sigma(fh, g) + \sigma(f, hg) - \sigma(f, h) = 0. \quad (29.5.4)$$

The is the defining equation of a 2-cocycle in the bar resolution, i.e. $d^2\sigma = 0$.

The second part of the theorem is showing that two different factor sets define the same extension iff they differ by a coboundary. If y_g is another system of preimages with a corresponding factor set τ , then $y_g = a_gx_g$ for some $a_g \in A$. We now write $y_hy_g = a_hx_ha_gx_g$ and expand both sides using the factors. This leads to

$$\tau(h, g)a_{hg} = a_h(ha_gh^{-1})\sigma(h, g), \quad (29.5.5)$$

which, after being put in additive notation, is exactly the statement that τ and σ differ by a coboundary.

It remains to check that this is an isomorphism of groups, which we omit (it is an easy exercise given that we have explicit equations for the bijection). ■

In summary, H^1 describes the conjugacy classes of embeddings of G into a semidirect product, whereas H^2 describes extesions with a given action that differ from a semidirect product (i.e. the semidirect product corresponds to the zero element of H^2 because its factors are everywhere equal to 1, i.e. the cocycle is trivial).

29.6 Relation to topology

Now we will consider an action of a (discrete) group G on a topological space X by homeomorphisms.

Proposition 29.6.1 ([27, Prop. 9.44]). *If a group G acts on a topological space X , then the singular chain complex $C_\bullet(X)$ is a complex of G -modules. Moreover, if the action of G is free, then each $C_n(X)$ is a free G -module.*

Proof. That $C_n(X)$ is a G -module is obvious. Also the action of G maps faces of a singular simplex to the corresponding faces of the image simplex, therefore ∂ is a G -equivariant map. Pick a subset $X_0 \subset X$ that overlaps each orbit of the G -action in exactly one point. Next one considers the set of all basic simplices, which are singular simplices whose 0-th vertex is

mapped into X_0 . Basic n -simplices generate all of $C_n(X)$ because each $\sigma \in C_n(X)$ can be written as $g(g^{-1}\sigma)$ for some g such that $g^{-1}\sigma$ is basic.

In the case of a free action it remains to check that the set of all basic n -simplices is a basis of $C_n(X)$ as a G -module: if $g\sigma_1 = h\sigma_2$ with σ_1, σ_2 basic, then $g = h$ and $\sigma_1 = \sigma_2$. Indeed, the former equality implies that the 0-th vertices of σ_1 and σ_2 coincide since they lie in the same orbit and at the same time in X_0 . But then, since the action is free, $g = h$, and since G acts by homeomorphisms, $\sigma_1 = \sigma_2$. Finally, if $\sum_j \alpha_j \sigma_j = 0$ with distinct σ_j and $\alpha_j = \sum_k m_{jk} g_k \in \mathbb{Z}[G]$ with distinct $g_k \in G$, then $\sum_{j,k} m_{jk} (g_k \sigma_j) = 0$. Hence all simplices $g_k \sigma_j$ are distinct, so each $m_{jk} = 0$ and thus $\alpha_j = 0$. ■

Definition 29.6.2 (Acyclic space). A space X is called acyclic if $H_0(X) \cong \mathbb{Z}$ and $H_n(X) = 0$ for all $n \geq 1$.

Theorem 29.6.3 ([27, Prop. 9.44]). *If a group G acts freely on an acyclic space X , then the singular chain complex $C_\bullet(X)$ is a deleted G -free resolution of the G -trivial module \mathbb{Z} .*

Proof. We know that $C_\bullet(X)$ is a complex of G -free modules. That X is acyclic sats that $\rightarrow C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \rightarrow 0$ is exact at each $n \geq 1$, hence $\rightarrow C_2(X) \xrightarrow{\partial_2} C_1(X) \rightarrow \text{coker } \partial_1 \rightarrow 0$ is exact. But, since X is acyclic, $\text{coker } \partial_1 = H_0(X) \cong \mathbb{Z}$. ■

Now let A be a G -module. $H^\bullet(X, A)$ is defined as the homology of the complex $\text{Hom}_G(C_\bullet(X), A)$. If X is acyclic and G acts freely, then $H^\bullet(G, A)$ is the homology of $\text{Hom}_G(S_\bullet(X), A)$ by Theorem 29.6.3.

Now suppose, in addition, that the action $G \curvearrowright X$ is properly discontinuous (in the context of topological actions this is just called proper). Then there is an isomorphism of complexes (see MacLane, “Homology”, pp. 135-136)

$$\text{Hom}_{\mathbb{Z}}(C_\bullet(X/G), A) \cong \text{Hom}_G(C_\bullet(X), A) \quad (29.6.1)$$

whenever A is G -trivial and X/G is the orbit space of X . But isomorphic complexes have isomorphic homology:

$$H^n(X/G, A) \cong H^n(G, A). \quad (29.6.2)$$

The next step is to exhibit a space X satisfying all the conditions above. First we state, without proof for now, the existence of a space whose fundamental group is G and all other homotopy groups vanish.

Definition 29.6.4 (Eilenberg-MacLane space). Given a group G , there exists an Eilenberg-MacLane space $K(G, 1)$ that is path-connected, aspherical ($\pi_n(K(G, 1)) = 0$ for $n \geq 1$), and whose fundamental group is $\pi_1(K(G, 1)) \cong G$.

The desired space X is then the following object.

Definition 29.6.5 (Classifying space). The universal covering space of $K(G, 1)$ is called the classifying space of G , denoted $\mathbf{B}G$. It is acyclic, G acts properly on $\mathbf{B}G$, and $\mathbf{B}G/G \cong K(G, 1)$.

It follows from (29.6.2) that if A is a G -trivial module, then

$$H^n(K(G, 1), A) \cong H^n(G, A); \quad (29.6.3)$$

the cohomology of an abstract group G coincides with the cohomology of its first Eilenberg-MacLane space.

30 Lie algebra (Co)Homology

Use Wiebel

31 Sheaf Cohomology

31.1 Presheaves

31.2 Sheaves

31.3 Resolutions of sheaves

31.4 Sheaf cohomology

32 Čech Cohomology

32.1 Orientability and cohomology

32.2 Čech cohomology and its applications

32.3 Čech vs. sheaf vs. de Rham cohomology

32.4 Čech vs. singular cohomology

33 Applications

33.1 Poincaré duality

33.2 Thom isomorphism

33.3 Intersection theory

Use Bott & Tu.

33.4 Thom class and degrees

33.5 Euler class and Lefschetz numbers

Part V

Lie Theory

Derive Lie's Third Theorem 12.5.6 from Theorem 11.3.10 by using the fact that every Lie algebra is a semidirect product of a solvable and semisimple algebras. This is done in [40, Rem. 9.5.12].

Part VI

Gauge Theory and Index Theorems

34 Fiber Sequences

Follow Miller

35 Classification of Principal Fiber Bundles

36 Characteristic Classes

37 Chern-Weil Theory

38 Clifford Algebras and Dirac Operators

39 Index Theorems

References

- [1] S.I. Gelfand and Y.I. Manin, “Homological algebra”, 2nd Ed., Springer, 2003.
- [2] B. Milewski, “Category Theory for Programmers”.
- [3] Jet Nestruev, “Smooth Manifolds and Observables”, Springer, 2003.
- [4] Connes and Marcolli, “Noncommutative Geometry, Quantum Fields and Motives”.
- [5] https://proofwiki.org/wiki/Topological_Product_of_Compact_Spaces.
- [6] F. Schuller’s video lectures on Differential Geometry and Physics.
- [7] J.M. Lee, “Introduction to Smooth Manifolds”, Springer, 2nd Ed., 2013.
- [8] J.M. Lee, “Introduction to Topological Manifolds”, Springer, 2nd Ed., 2011.
- [9] P.M. Gadea et al., “Analysis and Algebra on Differentiable Manifolds: A Workbook for Students and Teachers”, Springer, 2nd Ed., 2013.
- [10] J.M. Lie, “Riemannian manifolds”, Springer, 1st Ed., 1997.
- [11] J. Jost, “Riemannian Geometry and Geometric Analysis”, Springer, 6th Ed., 2011.
- [12] <https://mathoverflow.net/questions/47569/what-makes-four-dimensions-special>.
- [13] Marcelo A. Aguilar and Carlos Prieto, “Fiber Bundles”, 2010.
- [14] Bjørn Ian Dundas, “Differential Topology”, 2013.
- [15] Hopf fibration visualization, <https://www.youtube.com/watch?v=AKotMPGFJYk>.
- [16] G. Rudolph and M. Schmidt, “Differential Geometry and Mathematical Physics, Part I”, Springer, 2013.
- [17] G. Rudolph and M. Schmidt, “Differential Geometry and Mathematical Physics, Part II”, Springer, 2017.
- [18] M.F. Atiyah, “Vector Bundles over an Elliptic Curve”, Proc.London Math. Soc. 3 7, 1957, 414-452.
- [19] Glen E. Bredon, “Topology and Geometry”, 1993.
- [20] J. Milnor, “Morse Theory”, Princeton University Press, 1973.

- [21] J. R. Munkres, “Topology”.
- [22] A. Hatcher, “Algebraic Topology”, 2002.
- [23] A. Hatcher, “Vector Bundles and K-Theory”, draft.
- [24] D. Huybrechts, “Complex Geometry: An Introduction”, Springer, 2005.
- [25] R. Bott and L.W. Tu, “Differential Forms in Algebraic Topology”, Springer, 1982.
- [26] T. tom Dieck, “Algebraic Topology”, AMS, 2008.
- [27] J.J. Rotman, “An introduction to Homological Algebra”, 2nd Ed., Springer, 2009.
- [28] G.M. Bergman, “An Invitation to General Algebra and Universal Constructions”, 2020
(also published by Springer).
- [29] T. Leinster, “Basic Category Theory”, 2016.
- [30] P. Perrone, “Notes on Category Theory with examples from basic mathematics”, 2019.
- [31] J.P. May, “A Concise Course in Algebraic Topology”, 1999.
- [32] E.H. Spanier, “Algebraic Topology”. Springer, New York, 1994.
- [33] J.J. Duistermaat and J.A.C. Kolk, “Lie Groups”, Springer, Berlin, 2000.
- [34] K.-H. Neeb, “Infinite-Dimensional Lie Groups”, HAL, 2009.
- [35] F.-W. Warner, “Foundations of Differentiable Manifolds and Lie Groups”, Springer, 1971.
- [36] A.L. Onishchik and E.B. Vinberg, “Lie Groups and Lie Algebras I”, Springer, 1993.
- [37] M.R. Sepanski, “Compact Lie Groups”, Springer, 2007.
- [38] J.R. Munkres, “Elements of algebraic topology”, 1984, Taylor&Francis Group.
- [39] A.R. Shastri, “Basic algebraic topology”, 2014, CRC Press.
- [40] J. Hilgert, K.H. Neeb, “Structure and Geometry of Lie Groups”, Springer, 2010.
- [41] R.W. Sharpe, S.S. Chern, “Differential Geometry: Cartan’s Generalization of Klein’s Erlangen Program”, 1997.
- [42] A. Cap, J. Slovák, “Parabolic Geometries I”, 2009.

- [43] I. Kolář, P.W. Michor, J. Slovák, “Natural operations in differential geometry”, 1993.
- [44] M. Vákár, “Principal Bundles and Gauge Theories”, 2011.
- [45] M. del Hoyo, “[Complete connections on fiber bundles](#)”, *Indagationes Mathematicae* **(27 4)**, 2015.
- [46] S. Kobayashi, “Differential Geometry of Complex Vector Bundles”, Princeton University Press, Princeton, 1987.
- [47] M. Spivak, “A comprehensive introduction to differential geometry”, vol. 2, 1999.
- [48] J. Figueroa-O’Farrill, “[On the intrinsic torsion of spacetime structures](#)”, 2020.

Index

- CW*-complex, 120
G-Set, 99
G-bundle, 213
G-structure, 213, 446
G-torsor, 388
 \mathbb{R}^∞ , 125
 \mathbb{S}^∞ , 125
n-compressible space, 151
n-connected pair, 152
n-connected space, 152
n-equivalence, 152
- Abelian category, 585
Absolute parallelism, 556
Acyclic complex, 626
Adjoint bundle, 434
Adjoint functor, 28
Adjoint representation, 183
Affine group, 389
Affine space, 388
Algebra, 19
Almost complex manifold, 573
Annihilator, 311
Aspherical space, 138
Atlas
 equivalence of, 71
 smooth, 71
 topological, 71
Augmentation map, 627
- Baker-Campbell-Hausdorff formula, 347
Balanced product, 429
Barycentric coordinates, 614
Base space, 90
Basis of topology, 46
Betti numbers, 651
Bloch sphere, 211
Bott periodicity, 269
- Boundary of a topological space, 43
Boundary operator, 616
Bruhat decomposition, 175
Bundle morphism, 212
Bundle reduction, 213, 428, 442
- Canonical bundle, 577
Category, 10
 dual, 11
 of arrows, 14
 of diagrams, 14
 of group-valued maps, 16
 of pointed sets, 11
 of set pairs, 11
 of sets, 11
- Cauchy sequence, 52
Cell complex, 120
Cellular map, 132
Center, 196
Centralizer, 166
Characteristic class, 106, 230, 235
Characteristic distribution, 312
Chart, 69
Chern class, 273
Chern number, 238, 273
Christoffel form, 471
Christoffel symbols, 543
Clopen set, 48
Closed map, 50
Closed set, 43
Closure, 43
Clutching function, 237
Clutching functions, 213
Cocurvature, 465
Cocycle, 214
Codimension, 84
Coequalizer, 34
Cofinal set, 610

- Cofinality of good covers, 610
 Cohomology
 de Rham, 601
 of a cochain complex, 619
 with compact support, 606
 Coimage of a morphism, 586
 Cokernel, 585
 Colimit of a functor, 578
 Commutator of vector fields, 283
 Compact space, 50
 Complete metric space, 53
 Complete solution, 281
 Complex, 600
 of chains, 618
 of cochains, 618
 of simplicial chains, 616
 simplicial, 614
 Complex manifold, 573
 Complex vector bundle, 573
 Complexification, 572
 Cone, 152
 Conformal metric, 562
 Conformal structure, 553
 Connected space, 48
 Connected sum of manifolds, 280
 Connecting homomorphism, 598
 Connecting morphism, 599
 Connection, 94
 compatible, 496
 irreducible, 496
 linear, 517
 reducible, 496
 Continuous functor, 579
 Continuous map, 44
 Contractible complex, 626
 Contractible space, 56
 Coproduct, 32
 Coretraction, 20
 Coslice category, 24
 Cotangent bundle, 254
 Covering Homotopy Property, 92
 Covering map
 smooth, 90
 topological, 90
 Critical points/values, 75
 Crosscap, 67
 Curvature
 general, 470
 Curvature endomorphism, 522
 Deck transformations, 101
 Deformation
 retraction, 56
 Degree, 93, 156, 637
 Local, 637
 Dense subset, 43
 Derivation, 78, 189
 Derivation on an algebra, 281
 Derivatives on locally convex spaces, 204
 Determinant bundle, 298
 Determinant of a linear map, 296
 Diffeomorphism, 73
 Differentiable structure, 71
 Differential form, 299
 Differential operator on the tensor algebra,
 285
 Dirac monopole, 494
 Direct limit, 124
 Directed set, 609
 Displacement form, 536
 Divergence, 306
 Dolbeault cohomology, 610
 Dolbeault operators, 576
 Dynkin's formula, 363
 Ehresmann fibration theorem, 211
 Eilenberg-MacLane space, 139, 704
 Elementary divisors, 166

- Embedding
 smooth, 83
 topological, 82
 Endomorphism, 10
 Enriched category, 38
 Equalizer, 34
 Equivalence of categories, 39
 Equivariant maps, 99, 179
 Euler characteristic, 651
 Exact functor, 589
 Exact sequence, 588
 short, 588
 Exhaustion, 74
 Exponential map, 342
 of a linear connection, 540
 Exponential mapping, 568
 Extension lemma, 74
 Exterior algebra, 299
 Exterior derivative, 300
 Exterior form, 296
 Exterior product, 297
 Fiber, 90
 Fiber bundle
 smooth, 209
 topological, 209
 Fiber subbundle, 233
 Fibered coproduct, *see* Pushout
 Fibered product, *see* Pullback (in category theory)
 Fibration, 92
 Hurewicz, 140
 Serre, 145
 Field, 18
 Final locally convex topology, 203
 Flat
 G -structure, 227
 fiber bundle, 227
 Flat chart, 291
 Flat module, 592
 Flow box, 274
 Flow domain, 274
 Flow of a vector field, 274
 Flow-out, 279
 Foliation, 294
 Fréchet manifold, 208
 Fréchet space, 201
 Frame bundle, 432
 Frame of a vector bundle, 245
 Free group, 24
 Functor, 11
 hom-, 13
 Fundamental group, 57
 Fundamental groupoid, 65
 Gauge group, 441
 Gauge transformation, 441
 Geodesic, 537, 567
 Good cover, 609
 Good pair, 634
 Gradient operator, 565
 Grassmann algebra, *see* Exterior algebra
 Grassmann manifolds, 409
 Grassmannian, 249
 Grassmannian manifolds, 435
 Group
 Euclidean, 181
 Group bundle, 225
 Group character, 323
 Groupoid, 65
 Haar measure, 323
 Half-density, 434
 Heisenberg group, 354
 Holomorphic differential form, 577
 Holomorphic tangent bundle, 574
 Holomorphic vector bundle, 243, 573
 Holonomic frame, 289

- Holonomy, 94
- Holonomy bundle, 499
- Holonomy group, 474
- Homeomorphism, 44
- Homogeneous space, 388
- Homological algebra, 260
- Homology
 - Local, 636
 - of a chain complex, 619
 - reduced, 628
 - simplicial, 616
 - singular, 627
- Homotopy, 55
- Homotopy equivalence, 57, 626
- Homotopy Extension Property, 127
- Homotopy groups, 117
 - relative, 118
- Homotopy Lifting Property, 92
- Hopf fibration, 211
- Horizontal form, 474
- Ideal, 195
- Image of a morphism, 586
- Immersion
 - smooth, 81
 - topological, 82
- Induced representation, 458
- Initial object, 23
- Injective object, 592
- Inner product, 301
- Integrability, 229
- Integrable distribution, 290
- Integral curve, 274
- Integral manifold, 290
- Integrating factor, 315
- Interior, 43
- Invariant factors, 165
- Invariant vector field, 277
- Involutive distribution, 290
- Isometry, 52, 561
- Isotropy group, 393
- Isotropy representation, 394
- Iwasawa decomposition, 174
- K-group, 269
- Kähler structure, 221
- KAN decomposition, 174
- Kernel, 585
- Killing equation, 565
- Killing vector field, 396, 565
- Klein geometry, 388, 448
- Knot
 - torus, 68
- Koszul calculus, 517
- Laplace-Beltrami operator, 565
- Lemma
 - compression, 131
 - Escape, 277
 - Flow-out, 279
 - Lebesgue, 91
 - Uniform time, 276
 - Urysohn, 54
 - Yoneda, 40
- Lens spaces, 68, 138, 139, 157, 240
- Levi-Civita connection, 552, 554
- Levi-Civita tensor, 564
- Lie algebra, 189
- Lie derivative of tensor fields, 286
- Lie derivative of vector fields, 283
- Lie derivatives
 - of functions, 253
- Lie group
 - locally convex, 208
 - representation, 182
- Lie group action, 178
- Lie ring, 18
- Lie transport, 288

Lift of a path, 91
 Limit
 direct, 582
 inverse, 582
 Limit of a functor, 578
 Local geodesic chart, 546
 Locally convex Lie group, 208
 Locally convex space, 200
 Locally Euclidean Space, 69
 Logarithmic derivative, 330
 Long exact sequence
 of homotopy groups, 119
 Loop, 57
 Loop space, 60
 Mackey complete space, 208
 Magnetic monopole, 494
 Manifold
 smooth, 71
 topological, 69
 with boundaries/corners, 72
 Mapping cylinder, 128
 Maurer-Cartan equation, 320, 322
 Mayer-Vietoris sequence, 603
 Metric, 51
 equivalence of, 52
 on a bundle, 560
 Metric connection, 554
 Metric space, 51
 Mittag-Leffler condition, 595
 Modular quasicharacter, 323
 Module, 18
 Moduli space, 511
 Monodromy, 93, 102, 474
 Monodromy representation, 333
 Monopole
 bundles, 454
 Moore space, 653
 Morphism, 10
 bi, 20
 epi, 20
 iso, 21
 mono, 20
 split epi, 20
 split mono, 20
 Natural isomorphism, 39
 Natural transformation, 35
 Nerve of a covering, 614
 Nilpotent
 Lie algebra, 364
 Lie group, 364
 Nondegenerate basepoint, 66
 Norm, 53
 equivalence of, 53
 Normal bundle, 263
 Normalizer, 196
 Normalizer of a subgroup, 102
 Normed vector space, 53
 Open map, 50
 Orbit map, 180
 Orbit type, 394
 Orientable vector bundle, 224
 Orientation, 551
 bundle, 304
 of a bundle, 304
 of a manifold, 304
 on a vector space, 303
 on simplicial complexes, 617
 Orientation-preserving maps, 304
 Overcategory, 24
 Paracompact space, 69
 Parallelism
 absolute, 224, 252, 468
 Parallelizable manifold, 252
 Partition of unity, 74

- Path, 49
- Path Lifting Property, 91
- Path lifting property, 91
- Path-connected space, 49
- Pfaffian system, 311
- Picard group, 273
- Pin groups, 173
- Poincaré lemma with compact support, 607
- Prüfer group, 684
- Principal fiber bundle, 225, 408, 425
- Principal homogeneous space, 388
- Product (in a category), 32
- Product (in topology), 47
- Projective object, 592
- Projective plane, 67
- Proper map, 50
- properly discontinuous action, 107
- Pseudo-Riemannian structure, 551
- Pullback (in a category), 33
- Pullback (of a tensor), 261
- Pullback bundle, 230
- Push-forward, 78, 261
- Pushout, 34
- Quasi-isomorphism, 626
- Qubit, 211
- Quotient map, 46
 - in smooth category, 89
- Quotient object, 584
- Rank of a map, 75
- Rank of abelian group, 165
- Rank theorem, 80
 - global, 81
- Realification, 572
- Regular points/values, 75
- Relative homotopy, 92
- Relative homotopy groups, 118
- Restriction of a fiber bundle, 232
- Retraction, 20, 56
- Riemann normal coordinates, 568
- Riemannian manifold, 561
- Ring, 17
- Ring module, 18
- Rolling without slipping, 536
- Scalar density, 434
- Second countable space, 69
- Section, 20, 87
- Section of a fiber bundle, 232
- Semidirect product, 180, 197
- Seminorm, 200
- Separable space, 43
- Separating family, 200
- Separation axioms, 54
- Sequence
 - Mayer-Vietoris, 603
- Simplex, 614
 - singular, 627
- Simply connected space, 58
- Skeleton, 38
- Smash product, 48
- Smooth map, 72
- Smooth structure, 71
- Snake lemma, 598
- Soldering form, 529
- Spencer differential, 550
- Spin groups, 173
- Split sequence, 588
- Stabilizer, 99, 393
- Stable isomorphism of vector bundles, 268
- Standard vector field, 537
- Stiefel manifolds, 389, 435
- Structural equation, 331
- Structure group, 213
- Sub-object, 584
- Submanifold, 84
 - embedded, 84

- immersed, 84
- initial, 84
- properly embedded, 84
- weakly embedded, 84
- Submersion
 - smooth, 81
 - topological, 88
- Subspace topology, 43
- Suspension, 58
- Syzygy, 656
- Tangent bundle, 250
- Tangent vector, 77
- Tautological 1-form, 530
- Tautological bundle, 249
- Tautological bundles, 414
- Taylor's formula, 289
- Tensor
 - bundles, 258
 - components, 258
 - on a vector space, 257
- Tensor algebra, 26
- Tensor product, 26
- Terminal object, 23
- Theorem
 - Ambrose-Singer, 500
 - Analytic Subgroup, 378
 - Arzelà-Ascoli, 55
 - Brouwer's fixed point, 56, 636
 - Cantor's intersection, 53
 - Cellular Approximation, 135
 - Closed Subgroup, 382
 - Closed subgroup, 368
 - Darboux, 558
 - Engel, 364
 - Equivariant rank, 180
 - Excision (homology), 633
 - Freudenthal's suspension, 156
 - Frobenius, 290, 313
- Global Frobenius, 294
- Heyne-Borel, 55
- Hopf-Rinow, 570
- Implicit mapping, 77
- Initial Subgroup, 384
- Inverse mapping, 77
- Level set, 85
- Lifting, 94
- Local Section, 88
- Monodromy, 93
- Newlander-Nirenberg, 575
- Nielsen-Schreier, 112
- Rank-nullity, 589
- Rectification, 289
- Rectification/Straightening, 280
- Sard, 75
- Seifert-van Kampen, 65
- Siefert-van Kampen, 63, 581
- Suspension, 635
- Tubular Neighborhood, 263
- Whitehead, 131
- Whitney approximation, 264
- Whitney Embedding, 262
- Willmore, 285
- Yamabe, 383
- Topological space, 42
- Topology
 - coarse, 45
 - final, 45
 - fine, 45
 - initial, 45
 - quotient, 45
- Torsion
 - intrinsic, 549
- Torsion coefficients, 650
- Torsion form, 530
- Torsion subgroup, 165
- Torsion submodule, 680
- Torus knots, 68

- Total space, 90
Transition functions, 213
Transport functor, 107, 467
Triangulation, 616
Trivial bundle, 210
Trivialization, 210
Tubular neighborhood, 263, 403
Twisted differential forms, 309
- Undercategory, 24
Unimodular group, 323
Universal cover, 97
Universal enveloping algebra, 31
Universal object, 23
Universal property
 of coproducts, 32
 of free groups, 25, 31
 of products, 32
 of pullbacks, 33
 of pushouts, 34
 of quotient maps, 45
 of tensor algebras, 28
 of tensor products, 27, 30
 of universal covers, 99
 of universal enveloping algebras, 31
- Vector, 77
Vector bundle, 242
Vector field, 252
Vector space, 19
Vector subbundle, 247
Vertical bundle, 465
Volume form, 305, 551
- Weak equivalence, 152
Weak topology, 120
Weakly contractible space, 133
Wedge product of differential forms, 297
Wedge sum, 48